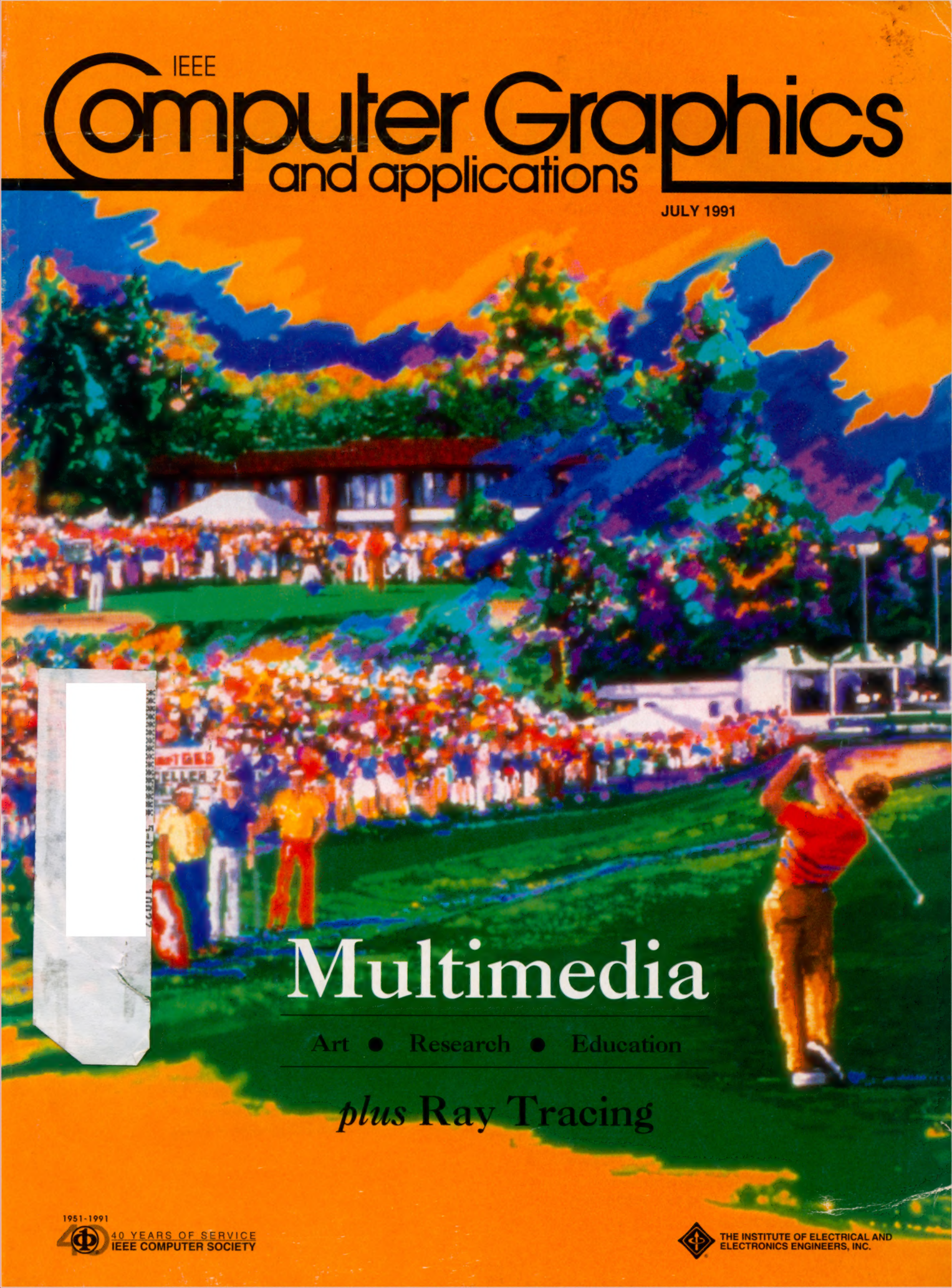


IEEE Computer Graphics and applications

JULY 1991



Multimedia

Art • Research • Education

plus Ray Tracing

1951-1991

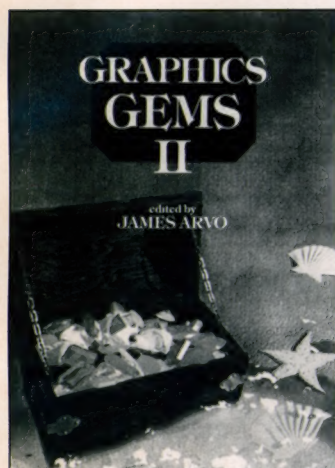


40 YEARS OF SERVICE
IEEE COMPUTER SOCIETY



THE INSTITUTE OF ELECTRICAL AND
ELECTRONICS ENGINEERS, INC.

Tricks of the Trade



Graphics Gems

edited by
Andrew S. Glassner

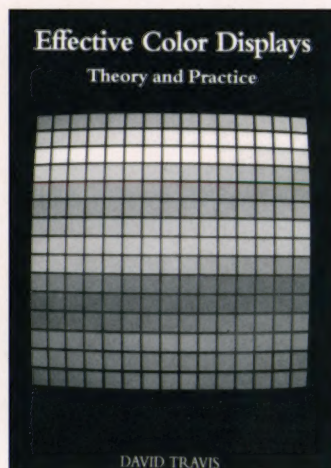
1990, 833 pp., \$49.95
ISBN: 0-12-286165-5

Graphics Gems II

edited by
James Arvo

July 1991, c. 536 pp., \$49.95 (tentative)
ISBN: 0-12-064480-0

Graphics Gems compiles the "tricks of the trade" in computer graphics, for example: an optimized inner loop, an essential snippet of geometry, an algebraic simplification that improves performance. **Graphics Gems II** offers more than seventy new "gems"—computer graphics programming techniques contributed by professional graphics programmers—covering ray tracing, rendering, image processing, curves and surfaces, and a new section on radiosity. Many of the gems contain public domain implementations in the C language which are complete and ready to run. Every programmer will find these handbooks essential resources for saving time and energy in daily programming. Buy both volumes for a toolbox full of tricks!



Effective Color Displays

Theory and Practice

David Travis

This book provides both the theoretical background and the practical guidelines for effective use of color in computer display. It may be used as a tutorial text for courses on color displays or as a practical guide for hands-on design. **Effective Color Displays** is a must read for programmers, engineers, and psychologists concerned with color applications at the user interface, with color illustrations, functions for color manipulation in C, look-up tables for color coordinates, and a checklist for display environments.

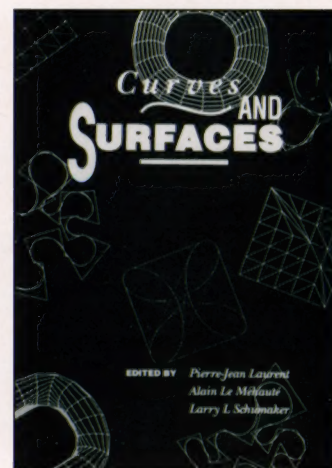
July 1991, 328 pp., \$49.95
ISBN: 0-12-697690-2

Controversial reading... Computerization and Controversy

Value Choices and
Social Changes

edited by
Charles Dunlop and Rob Kling

March 1991, 776 pp., \$34.95
ISBN: 0-12-224356-0



Curves and Surfaces

edited by
**Pierre-Jean Laurent,
Alain Le Mehaute, and
Larry L. Schumaker**

International contributors survey current topics such as wavelets, fractals, and 3-D reconstruction.

July 1991, c. 536 pp., \$49.95
ISBN: 0-12-438660-1

Oriented Projective Geometry

Jorge Stolfi

This book combines the elegance of classical projective geometry with the ability to talk about oriented lines and planes, signed angles, line segments, and convex figures. With this book, graphics programmers can assemble a rich, consistent, and effective set of tools for computational geometry that can be used in everyday work.

August 1991, c. 246 pp., \$29.95 (tentative)
ISBN: 0-12-672025-8

A Classic in its Field... Fractals Everywhere

Michael Barnsley

1988, 394 pp., \$44.50
ISBN: 0-12-079062-9

Order from your local bookseller or directly from



ACADEMIC PRESS Harcourt Brace Jovanovich, Publishers

Book Marketing Department #08071, 1250 Sixth Avenue, San Diego, CA 92101

Quote this reference number for free postage and handling on your prepaid order — 08071

Prices subject to change without notice. © 1991 by Academic Press, Inc. All Rights Reserved. LH/SS — 08071.

Reader Service Number 1

CALL TOLL FREE

1-800-321-5068

FAX **1-800-235-0256**

IEEE Computer Graphics and applications

IEEE Computer Society
10662 Los Vaqueros Circle
PO Box 3014
Los Alamitos, CA 90720-1264

Editor in Chief

Peter R. Wilson, Design Research Center, CII 7015
Rensselaer Polytechnic Institute, Troy, NY 12180

Submissions: Please submit six copies of all articles and proposals for special issues to Peter R. Wilson, editor in chief. All submissions are subject to editing for style, clarity, and space considerations.

Editorial Board

Ron Baecker	U. of Toronto
Frank W. Bliss	EDS/GM Corp.
Jack E. Bresenham	Winthrop College
Charles A. Csuri	Ohio State U.
Thomas A. DeFanti	U. of Illinois at Chicago
John C. Dill	Simon Fraser Univ.
Louis J. Doctor	consultant
Jose L. Encarnacao	Technical U. of Darmstadt
Nick England	Sun Microsystems
Andrew Glassner	Xerox PARC
Bertram Herzog	U. of Michigan
Tosiyasu L. Kunii	U. of Tokyo
Carl Machover	Machover Associates
Gregory M. Nielson	Arizona State U.
Michael J. Potel	Apple Computer
Michael L. Rhodes	Toshiba America MRI
Phillip K. Robertson	CSIRO, Australia
Lawrence J. Rosenblum	Naval Research Lab.
Jane Wilhelms	U. of California, Santa Cruz

Editor and Publisher: True Seaborn

Assistant Publisher: Douglas L. Combs

Editorial Director, Magazines: Marilyn Potes

Managing Editor: Nancy Hays

Staff Editor: Michael Haggerty

Membership/Circulation Mgr.: Christina Champion

Advertising Coordinators: Heidi Rex, Marian Tibayan

Design: Vic Grenrock

Production Artist: Kym Anjanette Williams

Art Consultants: Design & Direction

Assistant to Publisher: Pat Paulsen

Editorial Secretary: Karen Potes

Publications Board

Ronald G. Hoelzeman (chair), James Aylor, Victor R. Basili,
Richard Burke, Jon T. Butler, J.T. Cain, B. Chandrasekaran,
Carl Chang, Manuel D'Abreu, Dante Del Corso, Gerald Engel,
Michael Evangelist, James J. Farrell III, Tse-yun Feng,
Anil K. Jain, Ted Lewis, Ray Miller, Michael C. Mulder,
C.V. Ramamoorthy, True Seaborn, Sallie Sheppard, Harold Stone,
Earl E. Swartzlander, Peter R. Wilson

Magazine Advisory Committee

James J. Farrell III(chair), Fiorenza Albert-Howard,
Jon T. Butler, B. Chandrasekaran, Carl Chang, Manuel D'Abreu,
Dante Del Corso, Gerald Engel, Michael Evangelist,
Sushil Jajodia, True Seaborn, Pradip K. Srimani, Peter R. Wilson

Editorial: Unless otherwise stated, bylined articles and departments, as well as descriptions of products and services, reflect the author's or firm's opinion; inclusion in this publication does not necessarily constitute endorsement by the IEEE or the Computer Society.

Copyright and reprint permission: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limits of US copyright law for private use of patrons those articles that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 29 Congress St., Salem, MA 01970. Instructors are permitted to photocopy isolated articles for noncommercial classroom use without fee. For other copying, reprint, or republication permission, write to Permissions Editor, *IEEE Computer Graphics and Applications*, 10662 Los Vaqueros Circle, PO Box 3014, Los Alamitos, CA 90720-1264. All rights reserved. Copyright©1991 by the Institute of Electrical and Electronics Engineers, Inc.

Circulation: *IEEE Computer Graphics and Applications* (ISSN 0272-1716) is published bimonthly by the IEEE Computer Society, 10662 Los Vaqueros Circle, PO Box 3014, Los Alamitos, CA 90720-1264; (714) 821-8380. IEEE Computer Society Headquarters: 1730 Massachusetts Ave., Washington, DC 20036-1903. IEEE Headquarters: 345 East 47th St., New York, NY 10017-2394. Annual subscription: \$24 in addition to any IEEE group or society dues. Members of other technical organizations: \$39. Nonmember subscription rates are available on request. Back issues: members, \$10; nonmembers, \$20. This journal is also available in microfiche form.

Postmaster: Send address changes and undelivered copies to *IEEE Computer Graphics and Applications*, IEEE Computer Society, 10662 Los Vaqueros Circle, PO Box 3014, Los Alamitos, CA 90720-1264. Second class postage is paid at New York, NY, and at additional mailing offices.

Moving?

PLEASE NOTIFY
US 4 WEEKS
IN ADVANCE

MAIL TO
IEEE Computer Society
10662 Los Vaqueros Circle
PO Box 3014
Los Alamitos, CA 90720-1264

- This notice of address change will apply to all IEEE publications to which you subscribe.
- List new address below.
- If you have a question about your subscription, place label here and clip this form to your letter.

ATTACH
LABEL
HERE

Name (please print)

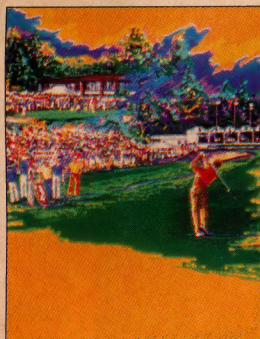
New address

City

State/Country

ZIP

IEEE Computer Graphics and applications



Multimedia artist Joni Carter has been painting sports images with the computer for nearly 10 years. She painted this bright cover for the Greater Greensboro Open. In August, nearly two million people will be able to buy Carter's artwork—on U.S. stamps. See About the Cover on page 5.

Cover art by
Joni Carter

FEATURE ARTICLE

14 Pixel-Selected Ray Tracing

Takaaki Akimoto, Kenji Mase, and Yasuhito Suenaga

Pixel-selected ray tracing is two to nine times faster than standard ray tracing. Its adaptive undersampling algorithm ensures that parts of thin or wedge-shaped objects are not lost.

MULTIMEDIA

24 Guest Editors' Introduction Multimedia—It's Actually Useful!

Jack Grimes and Mike Potel

26 Multimedia Scouting

Kristina Hooper Woolsey

These Multimedia Lab projects show how uninformed viewers can get interactive access to elements formerly reserved for graphics, film, and music professionals.

39 Numbers—A Medium That Counts

Glorianna Davenport and Jonathan D. Harber

As computational carriers of numbers, spreadsheets bring hypermedia to the office. A spreadsheet we developed helped us construct a template for the first hypermedia version of business case study.

45 Hypermedia and the Study of Ancient Culture

Gregory Crane

Allowing fluid movement among disparate data, hypermedia databases ideally suit the study of ancient cultures, where evidence ranges from literary texts to archaeological sites to individual objects.

52 Developing Multimedia Applications

Sorel Reisman

Experience with videodisc-based multimedia projects can provide development skill sets that you can use to produce your own multimedia applications. This multiphased approach shows you how.

58 Learning through Multimedia

Roy D. Pea

Teaching both new subject matter and new technology, Multimedia Works encourages students to create multimedia compositions that explain, tell a story, or persuade.

67 Cinematic Primitives for Multimedia

Glorianna Davenport, Thomas Aguiere Smith, and Natalio Pincever

Shots, sounds, and sequences advance the storytelling in film. Stratification helps us use these elements in multimedia.

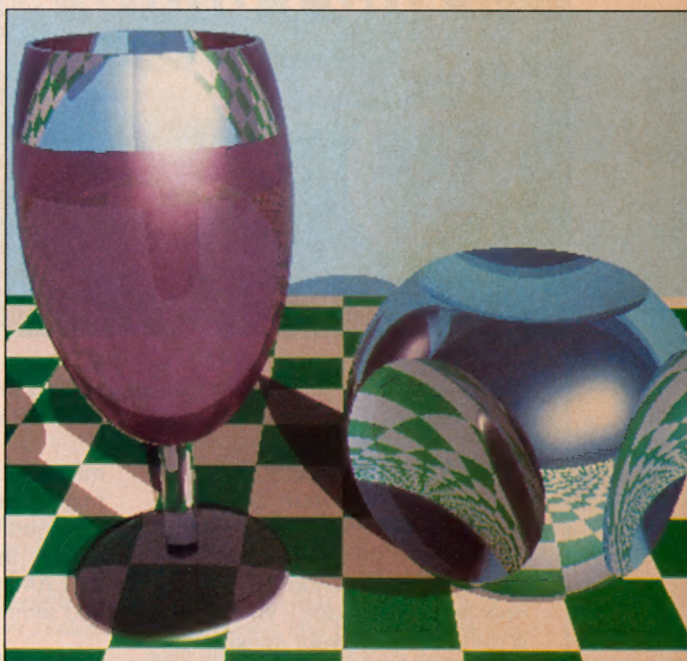
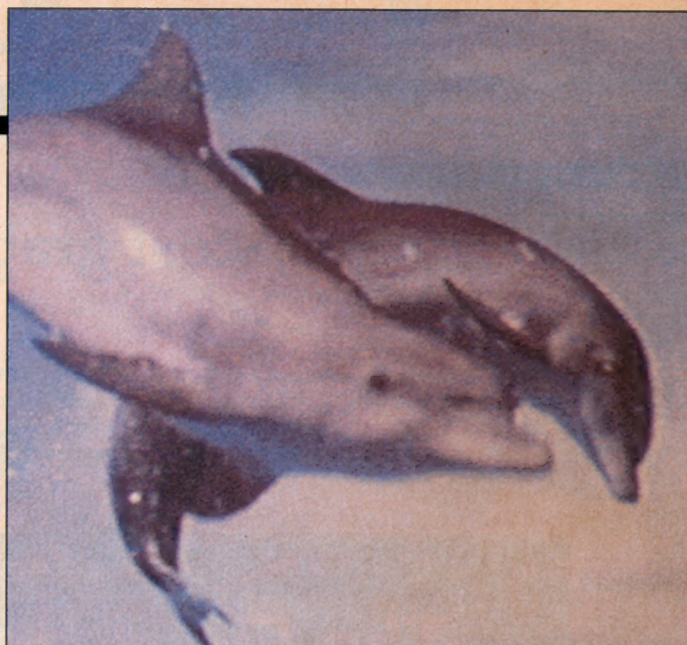
75 The Demise of Multimedia

David R. Clark

Forget "multimedia." We can best serve the future of interactive electronic presentations (IEPs) by developing special-purpose delivery engines rather than general-purpose computers.

DEPARTMENTS

- 4 **Letters**
Low-Cost Graphics Software for PCs?
- 5 **About the Cover**
Stamps of Approval
- 8 **Displays on Display**
Siggraph 91: Designs for Living
- 12 **Visualization Blackboard**
Acoustic Holography
- 81 **Jim Blinn's Corner**
A Trip down the Graphics Pipeline:
Pixel Coordinates
- 86 **Selective Update**
CG&A Seeks New Logo, NCGA 91,
Euro-DAC, Idaho Lab Looks for Art
- 91 **Book Reviews**
The Best Book on PHIGS
- 93 **New Products**
- 96 **Advertiser/Product Index**
- CIII **Final Focus**



Letters

Low-cost graphics software for PCs?

In the January 1991 issue of *CG&A* there is an article by Charles Csuri ("Art and Animation," part of the *CG&A* Tenth Anniversary Issue—ed.).

He states that there are free (Thompson Digital Images and Wavefront) software for independent artists. Do you have the addresses of these and any

other low-cost art (not animation) software for the IBM PC (under MS-DOS)?

Robert G. Thomas
Plymouth, England

Charles Csuri responds:

I keep forgetting that there is a PC world out there. My daughter, who is a computer graphics artist with a Mac, keeps telling me that there are better commercial packages than what I am using (a proprietary software), which is poorly documented and full of idiosyncracies.

Robert Thomas asks about software for the IBM PC, which presents difficulties, since I know very little about the PC. I really do not even know the software available for Macs, but my daughter is a good source to tell me about some of the packages and what they cost. Also, as an academician, I was not thinking of independent artists when I wrote my article.

When I made my comments about free software, I had in mind that if the artists are with an academic institution with a graphics workstation, APE (originally, Animation Production Environment) would be available, practically for free. At the time it would also have been available to independent artists. Now Ohio State University is licensing a company to sell and maintain the software. I guess it has proven to be too popular, and so, the university wants to make money. I know that software packages such as Wavefront, Alias, and Thompson Digital Images have been given by those companies to academic institutions for free on a case by case basis.

It would seem that Thomas would not be asking the question in the first place if there were low-cost graphics packages for his computer. The IBM PC may be a good computer, but perhaps a Mac would open the door to some excellent, relatively low-cost software.

Charles Csuri
Advanced Computing Center
for the Arts and Design



Break All The Traditional Barriers With Three-Dimensional Digitizers

From Science Accessories

Capture graphics from any three-dimensional object or physical movement within a 9 ft. x 9 ft. x 9 ft. cube. What you choose to input is only limited by space... your output is only limited by your imagination.

With a GP-8-3D sonic digitizer from SAC, you can input data directly from any three-dimensional object, or motion study. Your output can be flat art, contour images or dimensions of solid objects, or databases for film and video productions, robotics and more.

Systems Compatibility

The SAC GP-8-3D sonic digitizer converts physical information into digital form for processing by host computers or recording and trans-

mission equipment. Processed information can be used to provide real-time, multipoint analysis of motion, robotic instructions, and databases for CAM, CIM, and other automated and semi-automated systems.

Versatility

The GP-8-3D supports many popular software packages designed for your specific application. If you have special requirements, SAC will work with you to customize a software driver to meet your specifications.

Digitizing Today... With Tomorrow's Technology

SAC SCIENCE
ACCESSORIES
CORPORATION

200 Watson Boulevard P.O. Box 587
Stratford, CT 06497
(203) 386-9978 FAX (203) 381-9270 Telex 964300

About the Cover

Stamps of Approval

Michael Haggerty
Staff Editor



© 1991 Joni Carter

Carter originally designed this image for IBM in Atlanta. She adapted it for the packaging of her Olympic stamps.

A multimedia artist has designed the first set of U.S. stamps ever painted on the computer. On July 12, to coincide with the opening of the U.S. Olympic Festival in Los Angeles, the U.S. Postal Service and artist Joni Carter will issue five stamps commemorating Olympic track and field events. Sports enthusiasts and Olympic hopefuls will be among the first to see Carter's computer-painted designs.

But the first official unveiling of the 29-cent stamps occurred more quietly at Comdex 91 in Atlanta on May 21. There Joni Carter displayed her bright-colored images of sprinters, hurdlers, javelin and discus throwers, and a high jumper, and she also discussed how she created them.

Working on what she calls her "traveling PS/2 studio"—video camera, Targa

Micro Channel Board, Wacom Tablet, pressure-sensitive pen, and two IBM PS/2s—Carter studied video stills, then sketched and painted the five sports images. She then printed them on a color video printer in 512-line resolution for proofs. Once the postal service approved these designs, she output them in a higher resolution (1,024 × 768 line). For the U.S. Postal Service's final round of approval, she output 4 × 5-inch and stamp-size copies. The U.S. Postal Service then digitized and engraved the images with a Scitex printing system.

Multimedia correspondent

These are the most recent examples of the hundreds of sports images that Carter has created over the past 10 years. If there's been a major sports event, chances are good that Carter has

been there. As a computer correspondent, she has covered the Kentucky Derby, Indianapolis 500, French Open, baseball's All-Star Game, and two Super Bowls. Her paintings have been seen on the pages of *USA Today* and on major television networks in the United States and Canada. With her previous PC (an AT-based system), she painted images of the Olympics in Los Angeles, Lake Placid, and Calgary. With her PS/2s, she plans to be on the sidelines for the 1992 Olympics in Albertville (France) and Barcelona as well. This August, after showing some of her artwork in a number of southern California galleries, she'll be unveiling the official Professional Golfers of America commemorative painting.

Carter has also done some non-sports work. She designed material for televi-



Depending heavily on multimedia technology, Carter designed these stamps in five strong colors to correspond with the Olympic rings. These mark one of the first instances of completely digital stamp production.

sion, creating paintings for *Fantasy Island* and *Hart to Hart*. She created computer animation for *Knight Rider* and music videos ("some of the first computer graphics animation for rock videos," she says). She has worked in film, creating the omnipresent video displays in *Star Trek IV* and set decoration for *48 Hours*, and she has also designed posters, brochures, and interactive marketing material.

The technology that Carter uses lets her be this busy and prolific. With M-Motion software, she can freeze and capture an image on video, then use it as a model to paint from using Lumena software. She can then combine her painting with music, narration, or more graphics. To compose a multimedia presentation, she uses Audio Visual Connection (AVC) software, a program that lets her edit the video and/or still images in a number of ways (with cross-fades, dissolves, and so forth).

End user at the beginning

Carter has not always been convinced of the computer's power. Over 15 years ago, unemployed and adrift in Mammoth, California, she began painting with traditional media: oils and acrylics.

On May 14 of this year, at a session on multimedia at the Image World confer-

ence in Los Angeles, Carter explained that she became an artist almost on a whim. "I'm what you call 'a self-taught artist,'" she said. One day she decided to paint images of the subject she knew and loved best—skiing. After she finished her first paintings, she called a friend and said, "I've got these paintings. Maybe you'd like to hang them in your restaurant." The day after he hung them, he called her back and said, "Guess what. They sold." This news was a surprise and a revelation to Carter. "Well, I know what I am," she said, "I'm an artist."

Carter continued to paint pictures of skiers, and they sold well. She began painting other sports—baseball, basketball, track and field—but she wanted even more excitement in her work. She wanted to be on the field with the athletes, painting where the action was. That's when she turned to computers. In 1982 she sat down in front of one and taught herself how to paint with it.

In the early eighties, painting on the computer was a new science, and she was in on it at the beginning. She knew Richard Shoup as he was working on color painting for Xerox. She had a contract with Activision, the small, now defunct company in Pasadena, California that worked on computer graphics.

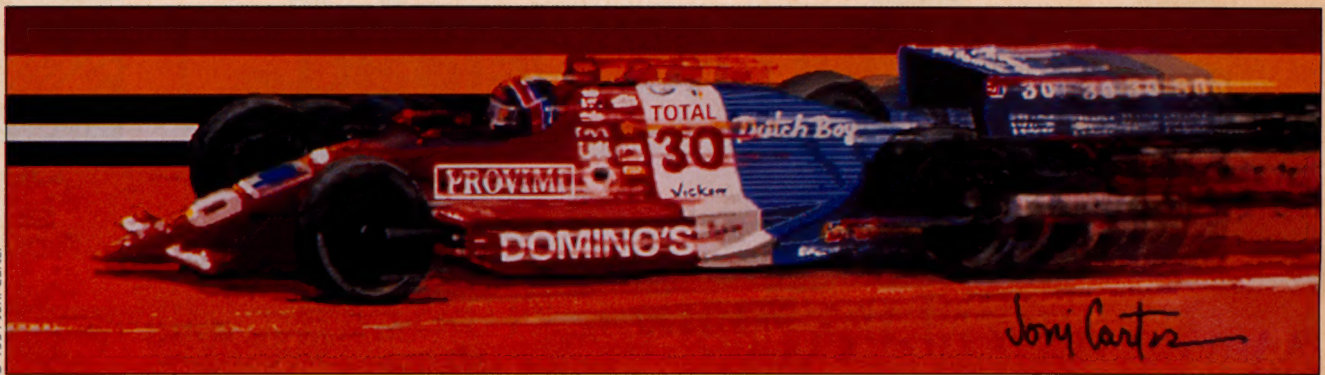
("I've worked with tons of companies that no longer exist," she said.) She remembers visiting George Lucas' early computer graphics laboratory when it was a modest workshop, and she recalls an early Siggraph conference in which Shoup and Alvy Ray Smith (another PC color pioneer) and she sat in a huge, mostly empty room in Los Angeles' Century City.

"Computer graphics was a small world to start with," Carter admitted in an interview after her Image World session. As a skilled software user, Carter had a unique place in that world as a beta tester for computer companies. And Carter brought a different perspective to computer graphics: Her primary interest was sports, she was a female, and she was nontechnical.

"I just like to paint"

Ten years and numerous companies and products later, she admits that she's still nontechnical. "I'm a techno-boob," she quipped during her session. "You can quote me on that." Later, during a glitch in her presentation (as frequently happens in multimedia sessions), she said, "Nobody said I was a technological genius. I'm not. I just like to paint."

Working in her Los Angeles studio, she depends on a group of software



Carter painted this image of one of the cars from the Indianapolis 500 from a still she obtained with video capture software.

users who test certain products, a researcher (Kate Richardson, who keeps up with hardware and software changes), and an engineer, Steven Cohn of Image Microsystems in Redondo Beach, California.

"She's just creative," Cohn said. An electrical engineer who has been working in multimedia since 1979, Cohn explained how he works with Carter. "She's an artist," he said. "She has these ideas and she brings them to me, and I help her make them work." For instance, Cohn wrote an addition to the M-Motion program that lets the system zoom up to full screen display—an obvious advantage for a computer artist.

Limits of multimedia

Like any experienced software user, however, Carter and her assistants have come up against some problems that they could not solve. Some of these failings, such as imprecise and vague user's manuals, seem easy to correct. But other problems, such as the awkward shapes of Lumena's brushes and the sluggishness of certain programs, will demand more development.

Perhaps the biggest problem Carter encounters is insufficient hard-copy output. "I'd give anything in the world if there was some standardization of color," she said. "The biggest tragedy of hard-copy output is that what's on your screen is what's on your screen, what's on their screen is what's on their screen, and never the two shall meet." This flaw in transferring files from system to system can result in slide services' producing slides with colors that differ distressingly from what the original creators saw on their screens. Standards for color output are still in flux. "I guess the biggest wish list is in hard-copy technology. That's the furthest behind."

While waiting for new standards and software improvements, Carter continues to take advantage of other innovations. She's now able to combine AVC and M-Motion, letting her edit still pictures, run full motion video, and make the graphics all at once. The integration of these technologies comes from the scientists she depends on—the software developers.

"There's a certain beauty in the science of development," Carter said. "A lot of software and hardware developers are real artists in their own right." But as Carter explained, these artists see things differently from their audience (the product users). In getting caught up in developing certain features, they sometimes overlook a basic fact: Maybe no one needs this feature, or, as Carter states it, "the applications might not be there." And Carter would know. She's admittedly application-driven. In the early days, she responded to graphics innovations (after the genuine and obligatory "wow") with "make it work for me. How can I apply it?"

Talent through technology

In the fall, Carter will begin applying her multimedia system to create an educational video for IBM. It will be targeted to kindergarten- through college-level students. In these videos Carter wants to teach people that computers are not monolithic, sterile objects. As Carter explains it, multimedia can be helpful in any field of study. She wants people to realize that the computer "is a tool, not a subject." How you use any tool depends on your skill and creativity.

"I used to send videotapes of my art to the networks," Carter began. "They'd send them back and say, 'Thank you very much, Joni, but we al-

ready have a [Quantel] Paintbox.' I said, 'When someone submits a script, do you say "Thank you very much, we already have a word processor?" Are you comparing me to the machine?' The network mistakenly thought that Carter simply took photographs and manipulated them. They confused her work with the computer. "The message I have to get through is that I still have to have the *talent* to go *through* the computer."

If her latest success and busy schedule are any proof, then Carter obviously has that talent. She'll soon be working on other projects with *USA Today* and the networks, and a plan to upgrade her system to an IBM RISC 6000 is also in the works. On July 13, 1.9 million copies of the stamps she created will be available across the United States. Because stamps eventually make their way into philatelists' collections, this is perhaps the largest distribution of collectible computer artwork to date. "Well," she said, "I know what I am. I'm an artist." □



Joni Carter with her PS/2 studio.

Non-Stop



destroy circuits

CSCA Guest Speaker

(non-stop design)

UDY

VANDERLAN'S

em'a grā)

m a g
a z i n
e

April 19
THURSDAY

6:30 Social hour
7:30 ANNOUNCE
7:45 present

INS REGARDING ATTY GEN DETAIL OF LONG ISLAND AREA-NUMBER CODE

TIME	PLACE	AREA-NUMBER CODE
9 1022AM TO TORONTO	ON 416 362-1913 D	1 6 % P
9 151PM TO BERKELEY	CA 415 845-9021 D	1 1
9 446PM TO NEW YORK	NY 212 807-4000 D	23 1
21 158PM TO BERKELEY	CA 415 845-9021 N	2 8
21 231PM TO BERKELEY	CA 415 845-9021 N	35 2
21 238PM TO BERKELEY	CA 415 845-9021 N	2 1
21 246PM TO BERKELEY	NY 212 279-2900 D	2 2
22 1034AM TO NEW YORK	NY 212 332-0400 D	2 1
22 1034AM TO LOS ANGELES	CA 805 295-1850 D	2 2
22 1034AM TO NEW YORK	CA 212 807-4000 D	2 2
22 1034AM TO NEW YORK	NY 212 362-1913 D	2 2

Displays on Display

Editor: Charles A. Csurí

Siggraph 91: Designs for Living

Michael Haggerty
Staff Editor

On the facing page is a detail from John Weber's *Design Circus* poster (originally 13 $\frac{1}{16}$ × 20 $\frac{1}{4}$ inches), a playful image from the Siggraph 91 Art and Design Show. Weber designed this poster on an Apple Macintosh IIfx for the Columbus Society of Communicating Arts. He used Aldus Freehand, Image Studio, Superpaint, and Pagemaker software.



Susan Sifton designed the graphics for the book *Symmetry* (9 × 8 inches) for the Los Angeles Contemporary Exhibition. She used Aldus Freehand and Pagemaker 4.0 on an Apple Macintosh IIfx.

Siggraph's usual art show is a thing of the past. It has gone the way of the direct view storage tube and the faded proceedings from Siggraph 78. This year the old art show has been replaced by the Siggraph 91 Art and Design Show, an exhibition of both fine art and design. From July 28 to August 2, attendees of the eighteenth annual international conference on computer graphics in Las Vegas, Nevada will be able to see innovative print, product, and architectural design as well as fine art.

Expanding the art show to include design is the brainchild of this year's art and design show chair, Isaac Victor Kerlow (chairman of the computer

graphics department at the Pratt Institute). He decided to expand the show to allow for the impressive submissions that are not exactly "fine arts."

As Kerlow explained in a recent interview, the idea to expand the show came to him during the judging of last year's show. "From looking at the submissions," he began, "it became obvious to me that they [the judges] were throwing out a lot of good projects because they

didn't fit the fine arts category." Some of these projects nevertheless used computer graphics attractively and skillfully.

Kerlow decided to enlarge the scope of the art show for artistic and economic reasons. "I wanted to do this for two reasons," he said. "The first is there is a lot of good work out there that hasn't been shown and talked about." And the second reason: "From the marketing point of view, it is silly to think that buyers of



This Intertrans Annual Report illustration (8½ × 11 inches) was designed by Linda Helton. Helton and computer artist Lisa Johnson used an Apple Macintosh IIx, dual page monitor, Laserwriter IINT, and Superpaint 2.0, Adobe Illustrator 1.9.3, and Quark Xpress software.



Carol Bouyoucos designed "Graphika" (book, 18½ × 12 inches) at Waters Design Associates for James River. She used Apple Macintosh IIcx and Lightspeed Design 10 hardware with Quark Xpress and Adobe Illustrator software.

the equipment are not designers but are only artists." Enlarging the art show to include design more accurately reflects the work actually being done in computer graphics. "We must confront the reality that computers are being used extensively in the graphic arts," Kerlow said.

Kerlow cited a number of statistics that support this realistic view of the graphics market. While only a few fine artists use the computer, most of the graphic arts (print, page, and product design) is computer-based. At some point in their production, almost every printed piece is now computerized.

For the first time, design materials will be sufficiently represented at Siggraph. These materials make up the 2D category of the design show and fall into one of a number of subcategories: poster, book, brochure, illustration, or stationery. The 3D category is also divided into broadly defined subcategories: product design, architectural rendering, and CAD/CAM.

Glimpses of a maturing art

Kerlow especially praised the 3D entries in both the fine art and design categories (3D images will be featured in the next issue of *CG&A*). "The quality of the 3D entries is excellent," Kerlow said. "3D computer graphics is finally a very mature set of techniques. It really works now for artists and designers."

Commenting further on the maturity of computer graphics technology, Kerlow noted that the overall quality of all submissions for both categories was commendable. Even though there were fewer entries than last year, the quality was just as good. The fine arts jury—Timothy Binkley, Eleanor Flomenhaft, and Cynthia Goodman—looked at 3,000 entries before choosing the 50 that will be shown. (Judson Rosebush joined the fine arts jury for the selection of computer animation.) Design jurors

Kent Hunter, David Peters, and Wendy Richmond narrowed 700 entries down to the 60 in the exhibit. (Donald Rorke assisted the design jury for the selection of entries in the 3D category.)

The systems artists are using also reflect the industry's maturation. Kerlow noted that as recently as three years ago many of the fine arts entries were done on VAXs. This year there was not a single VAX-based entry. About 70 percent of the fine art was done on Apple hardware; the rest was on workstations and PC compatibles. Out of the 60 accepted design pieces, 50 were done on Apple hardware. Three were done on IBM PCs, and the rest on workstations. As industry analysts agree, computer power is clearly becoming more accessible, more affordable, and more widely used.

The design show was supposed to include examples of computer-human in-



Margo Chase designed this album jacket (12 1/16 x 12 inches) for Atlantic Records' recording artists Escape Club. Chase used Adobe Illustrator 3.0 and Quark Xpress on an Apple Macintosh IIfx. Logos and type were output on a Linotronic and then separated and stripped traditionally.

© 1991 Earl Gee, Mark Anderson Design



Earl Gee from Mark Anderson Design created a series of three posters for Z Typography. Two of the series are shown here (both originally 15 x 20 inches). Gee designed "Set Type in Your Sleep" and "24 Hour Turn-around" on an Apple Macintosh IIfx with Quark Xpress, Adobe Illustrator, Aldus Freehand, and MacDraw software.

terface design (icons, screen design, and so forth), but only a couple of entries were submitted in this category. Kerlow was disappointed in the low turnout. He attributed it to software developers' reluctance to display their work in an industry troubled with copyright litigation.

Redesigning the conference

Other Siggraph events are geared to the unique needs of designers using computer graphics. Two day-long courses specifically address design issues. On Monday, July 29 "New Media Applications in Art and Design" will introduce attendees to multimedia, artificial reality, and other new interactive technologies. On Tuesday, July 30 "Graphic Design in the New Media Environment" will attempt to fill the gap left by the lack of interface designs in the art and design show. This course addresses designing interfaces for new media such as multimedia and virtual reality.

Four panel discussions also address design issues. "Graphic Design in the Nineties: New Roles, Options, and Definitions" challenges attendees to examine the relationship of creativity and the computer. "What Next? A Provocative

Look at Curriculum, Creativity, and Logistics in Teaching Artists and Designers to Use Computers" is designed for computer graphics educators. Aesthetic issues will be at the center of "Making Virtual, Artificial, or Real Computer Art," and multimedia and hypermedia will be the challenges of "Designing for New Media: Technologists and Visual Designers Working as a Team."

In addition to the technical sessions, courses, and panels, conference attendees will also be able to see virtual reality demonstrations in "Tomorrow's Realities," Siggraph's first juried exhibit of the new medium. A hypermedia exhibit is also planned, and, as usual, about 225 exhibitors will display their latest software, hardware, and peripherals. Also of interest to designers (and the graphics community as a whole) is the electronic theater, showcasing some of the most innovative animation, special effects, new techniques, and visualization.

This year, 25,000 people are expected to attend Siggraph, which is sponsored

by the Association for Computing Machinery's Special Interest Group on Computer Graphics in cooperation with the IEEE Computer Society's Technical Committee on Computer Graphics. Courses, sessions, panels, and other events are designed to help scientists, educators, students, and artists keep up with computer graphics' progress.

The Siggraph Art and Design Show and the conference's emphasis on design are intended to show off the changes in computer graphics. One of the most important changes is the wide acceptance computers now have among designers. The techniques for type manipulation, image processing, and 3D rendering have become integral tools for many designers. "Computer graphics is not an unstable, futuristic thing anymore," Kerlow said. Just five years ago, new technology was a novelty and the very idea of computer graphics was a joke in the design community, but now, as shown by the efforts of the artists on these pages, designers take computer graphics quite seriously. □

Visualization Blackboard

Editor: Lawrence J. Rosenblum

Acoustic Holography

Lawrence C. Schuette
Naval Research Laboratory

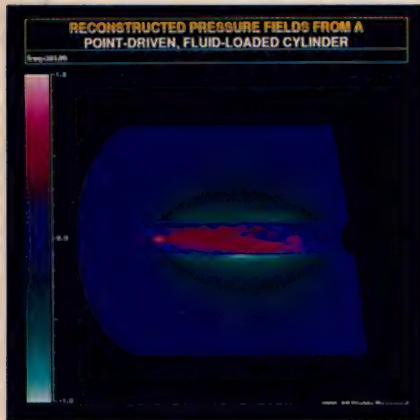


Figure 1. A volume of acoustic pressure surrounding a submerged cylindrical object at 391 Hz.

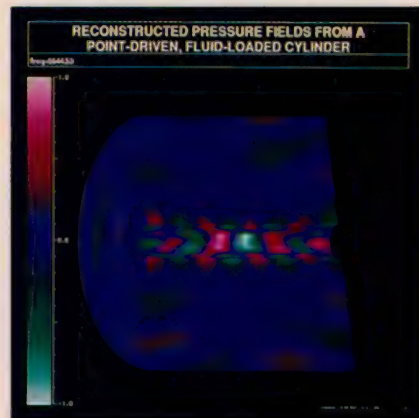


Figure 2. A volume of acoustic pressure surrounding a submerged cylindrical object at 5,645 Hz.

The technique of optical holography ("total recording") dates from the 1940s, when Dennis Gabor realized that both the phase and amplitude of light waves diffracted by an object could be captured on photographic film.¹ An optical holographic measurement, then, consists of two numbers: the light's intensity and its phase. More recently, researchers have extended holographic techniques from optics to acoustics. The acoustic analog to the optical technique consists of a monochromatic measure of the amplitude and phase of the acoustic pressure.

Acoustic holography has become an important diagnostic technique for analyzing radiating structures. To analyze a cylinder, we take a series of measurements—a scan—over a shell close to the surface of the cylindrical object. A computer-controlled robotic scanner at the

Naval Research Laboratory (NRL) provides the accuracy and repeatability necessary to create high-resolution holograms.²

A typical experiment to investigate a submerged, radiating, cylindrical object involves collecting a time-data record of 2,048 points at each of 4,096 equally spaced locations about the cylinder. Thus we collect pressure, $p(r_o, \phi_o, z_i, t)$, where r_o is the radial standoff distance, ϕ_o is the angular increment, z_i is the axial increment, and t is the time increment.

Once we have collected the data, we enter it into an analytical model and process it digitally. Inputs to this model include the geometry of the radiating object (in this case a cylinder) and the driving force exciting the object. The model's reconstructive techniques then produce three-dimensional information about the acoustic field exterior to the cylinder, the normal velocity of the cylinder's surface, the acoustic intensity radiated by the cylinder, the structural

intensity, and the mechanical energy flow. All of this information is available on a per-frequency basis from 300 Hz to 12 KHz.

Visualizing acoustic data

The collection and analysis of acoustic holographic data at NRL has centered around the Generalized Near Field Acoustical Holography (Genah) technique. Genah was developed to analyze submerged, radiating, cylindrical objects.³ The analysis of acoustic holograms generated with Genah has led to the development of a volume visualization and animation capability using a Silicon Graphics 4D Iris Graphics Workstation.

Because Genah results in a large data set (typically 32 radial, 64 angular, and 64 cells lengthwise for a total of 131,072 cells per frequency with 1,000 frequency

bins), scientific visualization techniques are essential. In the interest of interactivity, only the surfaces within view are drawn. The volume of data is rendered as six surfaces, so only a small fraction of the cells needs to be drawn. This allows us to manipulate the holographic volume quickly and gain insight into the nature of the acoustic fields.

We commonly visualize our results on a "per frequency" basis. Incrementing the frequency of the visualized portion of the data set lets us observe the frequency dependence of the acoustic pressure radiating from the cylinder. Figures 1 and 2 show a volume of pressure surrounding a cylinder at 391 Hz and 5,645 Hz, respectively. The red color indicates high positive pressure and the green, high negative pressure. The cylinder's response to the two frequencies differs noticeably.

Additionally, we can systematically enlarge the visible surfaces of the volume of pressure. By this technique we gain insight into the structure of the radiating pressure field for a given frequency. Figures 3 and 4 show an end view of the volume of pressure at 391 Hz. These pictures are slices of the pressure volume looking down the axis of the cylinder. Figure 3 shows the pressure field at the edge of the cylinder some distance from the exciting force, while Figure 4 shows the pressure field near the center of the cylinder, directly over the exciting force. Clearly, the pressure field is much stronger as it gets closer to the exciting force.

Finally, because a hologram stores a pressure phasor (amplitude and phase), we generate animations of the time varying pressure. We obtain this by multiplying each data point by $e^{j\omega t}$ and displaying the real portion of the resultant

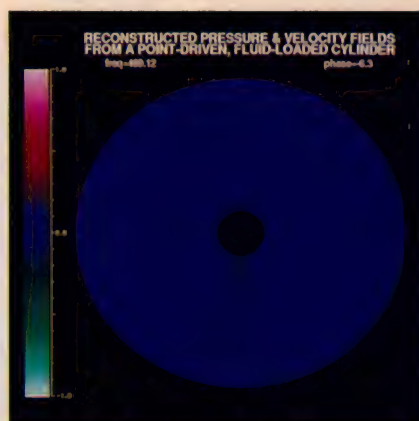


Figure 3. End view of acoustic pressure field at the edge of the cylindrical object, at 391 Hz, showing weak coupling to the driving force inside the cylinder.

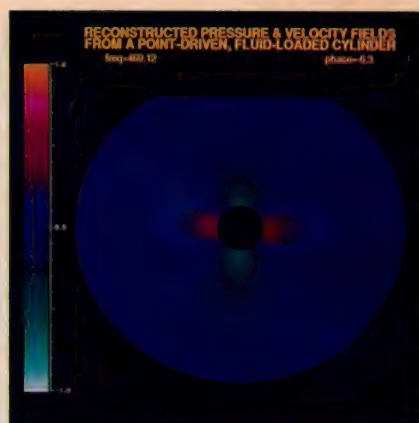


Figure 4. End view of acoustic pressure field at the center of the cylindrical object, at 391 Hz, showing strong coupling to the driving force inside the cylinder.

phasor. Smoothly varying t over one complete period then results in the animation of the time varying pressure field. Animation of the pressure field shown in Figures 5, 6, and 7 highlights the strong radiation that occurs from the edge of the cylinder at 5,645 Hz with t varying over one half period.

The application of volume visualization and animation techniques to acoustic holographic data has increased our understanding of the mechanisms of a radiating structure. A particularly intriguing phenomenon (the edge radiation mentioned previously) has become evident because of our ability to view the volume as a function of frequency and time.

Acknowledgments

I wish to acknowledge the efforts of Earl G. Williams and Karl B. Washburn for the Genah algorithm and experiments described.

References

1. D. Gabor, "A New Microscope Principle," *Nature*, Vol. 161, 1948, p. 777.
2. E.G. Williams, H.D. Dardy, and R.G. Fink, "Nearfield Acoustical Holography Using an Underwater, Automated Scanner," *J. Acoustical Society of America*, Vol. 78, No. 2, 1985, pp. 789-798.
3. E.G. Williams, H.D. Dardy, and K.B. Washburn, "Generalized Nearfield Acoustical Holography for Cylindrical Geometry: Theory and Experiment," *J. Acoustical Society of America*, Vol. 81, No. 2, 1987, pp. 389-407.

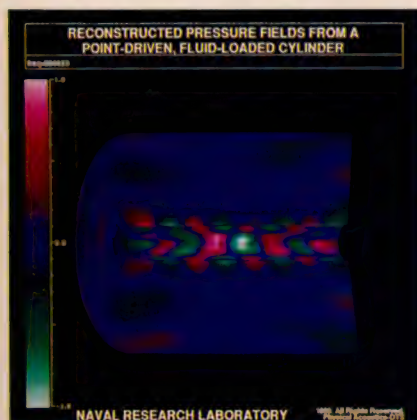


Figure 5. 0 degrees.

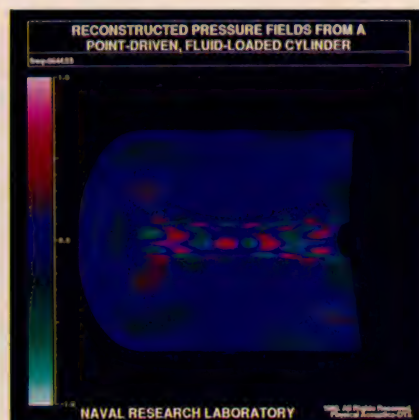


Figure 6. 90 degrees.

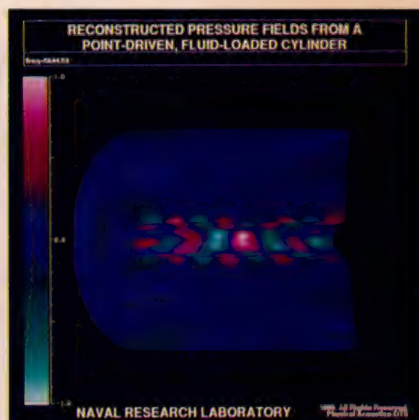
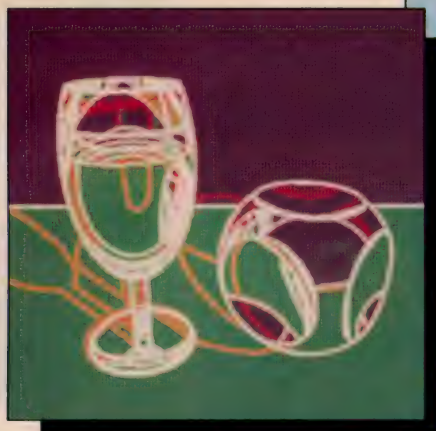


Figure 7. 180 degrees.

Figures 5, 6, and 7. The time varying acoustic pressure field at 5,645 Hz. This series shows the edge radiation while varying t over half a cycle.



Pixel-Selected Ray Tracing

Takaaki Akimoto, Kenji Mase, and Yasuhito Suenaga
Nippon Telegraph and Telephone

Pixel-selected ray tracing is two to nine times faster than standard ray tracing. Its adaptive undersampling algorithm ensures that parts of thin or wedge-shaped objects are not lost.

Ray tracing is a powerful technique for rendering reflection and refraction phenomena, but it requires excessive computer time. Various acceleration methods have been proposed to improve rendering speed (see sidebar). In this article we report on an acceleration method based on an idea that Whitted presented in the first paper on ray tracing.¹ He proposed making antialiased images by hierarchical adaptive oversampling, but we use hierarchical adaptive undersampling to reduce the number of pixels whose intensity must be calculated by ray tracing.

Our image-generation method, which we call pixel-selected ray tracing, takes advantage of the area coherence present in most pictures. PSRT cuts computing time by reducing the number of pixels for which intensities are calculated by regular ray tracing. We calculate the intensities of the other pixels by interpolation or by accelerated ray tracing using area coherence. Generally, undersampling causes such problems as wedge-shaped objects rendered without their tips. PSRT's efficient adaptive undersampling algorithm avoids such degradations.

Theoretical background for our use of area coherence came from the pencil tracing method proposed by Shinya, Takahashi, and Naitou.² They used calculations with the optical system matrix to estimate the paraxial rays and control the estimation accuracy around the axial ray. However, they did not discuss practical algorithms for adaptive undersampling and pixel-inten-

sity calculation. Lee, Redner, and Useton³ proposed statistically optimized sampling for distributed ray tracing and clarified the relationship between the number of rays and the intensity-value characteristics. Mitchell⁴ reported a technique for generating antialiased images using nonuniform sampling at low densities, but his purpose was to reduce oversampling for spatial antialiasing and motion blur, not to generate images faster with undersampling.

Pixel-selected ray tracing

Adjacent pixels are strongly correlated in typical images. In most images synthesized by ray tracing, adjacent pixels have very similar pixel intensities and ray-object intersection histories. PSRT uses this similarity to reduce the number of pixels that must be calculated by ray tracing. (In this article, we use *intensity* to express the color or brightness of a pixel or ray. For color images, the pixel intensity is a vector of the intensities for each color component.)

To implement PSRT, first we must find homogeneous regions in images. Because, strictly speaking, homogeneous regions are not completely detected until ray tracing calculates all the pixel intensities, we use hierarchical adaptive undersampling. We calculate the intensities of coarsely sampled pixels by regular ray tracing. Then, we examine the homogeneity of the small regions surrounded by the sampled pixels by checking the sim-

ilarity of the sampled pixels. In homogeneous regions, we calculate pixel intensity with a simple procedure such as interpolation. In nonhomogeneous regions, we calculate the intensities of pixels sampled at smaller intervals by regular ray tracing and repeat the procedure described above (the acceleration method using adaptive undersampling).

This adaptive undersampling method has several advantages. If an image-generation system has a routine for calculating pixel intensity by ray tracing, adaptive undersampling can be implemented for fast image generation. There are no limitations on the 3D model representations or the ray tracing procedures we can use it with, and we can combine it with methods that reduce ray-object intersection calculation. Adaptive undersampling is especially effective with larger images, because the homogeneous regions are larger.

But adaptive undersampling can result in some image-quality defects. Small objects and parts of thin or wedge-shaped objects may disappear when they are located between the initially sampled pixels. The intensity of the pixels representing the object or object part is interpolated using the intensity of the sampled pixels, and they do not have the same color as the object. While there are some methods to help detect small objects,⁵ it is extremely difficult to completely avoid having some small objects disappear when point sampling is used. On the other hand, we can ensure that parts of thin objects and tips of wedge-shaped objects do not disappear by improving the adaptive un-

a/s, so
-Whitted's
bounces
did this!

Ways to accelerate ray tracing

The *bounding volume* method, first presented in Whitted's 1980 paper on ray tracing,¹ speeds ray tracing by checking for ray intersections with bounding volumes around objects before performing the intersection calculation for each object. A hierarchy of bounding volumes reduces the calculation time even further.²

Glassner's *space subdivision* method accelerates the intersection calculation by dividing 3D object space into small rectangular solid regions, representing it with an octree structure, and extracting the regions through which rays pass.³ Fujimoto and Iwata⁴ proposed a modification: Partition 3D object space into uniform cubic regions or voxels, so a line of voxels indicates regions through which rays pass. Arvo's⁵ modification of the space subdivision method uses ray classification in a five-dimensional space composed of ray origins (three dimensions) and ray directions (two dimensions). These methods are extremely effective for a large number of objects when good space subdivision is possible. *Beam tracing*⁶ differs from other acceleration methods because it traces bundles of rays (beams) instead of tracing individual thin rays. However, beam tracing can handle only polygonal objects and cannot trace refracted beams correctly.

Proposals to speed up the generation of successive images include *incremental ray tracing*⁷ and *parameter-*

ized ray tracing.⁸ By reevaluating ray trees from the initial rendering, these methods allow quick redisplay of the same scene.

In this article, we propose *pixel-selected ray tracing*, a ray-tracing acceleration method based on hierarchical adaptive undersampling of pixels in the scene.

References

1. T. Whitted, "An Improved Model for Shaded Display," *Comm. ACM*, Vol. 23, No. 6, June 1980, pp. 343-349.
2. H. Weghorst, G. Hopper, and D. Greenberg, "Improved Computational Methods for Ray Tracing," *ACM Trans. Graphics*, Vol. 3, No. 1, Jan. 1984, pp. 52-69.
3. A.S. Glassner, "Space Subdivision for Fast Ray Tracing," *IEEE CG&A*, Vol. 10, No. 4, Apr. 1984, pp. 15-22.
4. A. Fujimoto, T. Tanaka, and K. Iwata, "ARTS: Accelerated Ray Tracing System," *IEEE Computer Graphics and Applications*, Vol. 6, No. 4, Apr. 1986, pp. 16-26.
5. J. Arvo, "Fast Ray Tracing by Ray Classification," *Computer Graphics (Proc. Siggraph)*, Vol. 21, No. 4, July 1987, pp. 55-64.
6. P. Heckbert and P. Hanrahan, "Beam Tracing Polygonal Objects," *Computer Graphics (Proc. Siggraph)*, Vol. 18, No. 3, July 1984, pp. 119-127.
7. K. Murakami, K. Hirota, and M. Ishii, "Incremental Ray Tracing," *Trans. Information Processing Soc. Japan*, Vol. 30, No. 6, June 1989, pp. 689-698 (in Japanese).
8. C.H. Sequin and E.K. Smyrl, "Parameterized Ray Tracing," *Computer Graphics (Proc. Siggraph)*, Vol. 23, No. 3, July 1989, pp. 307-314.

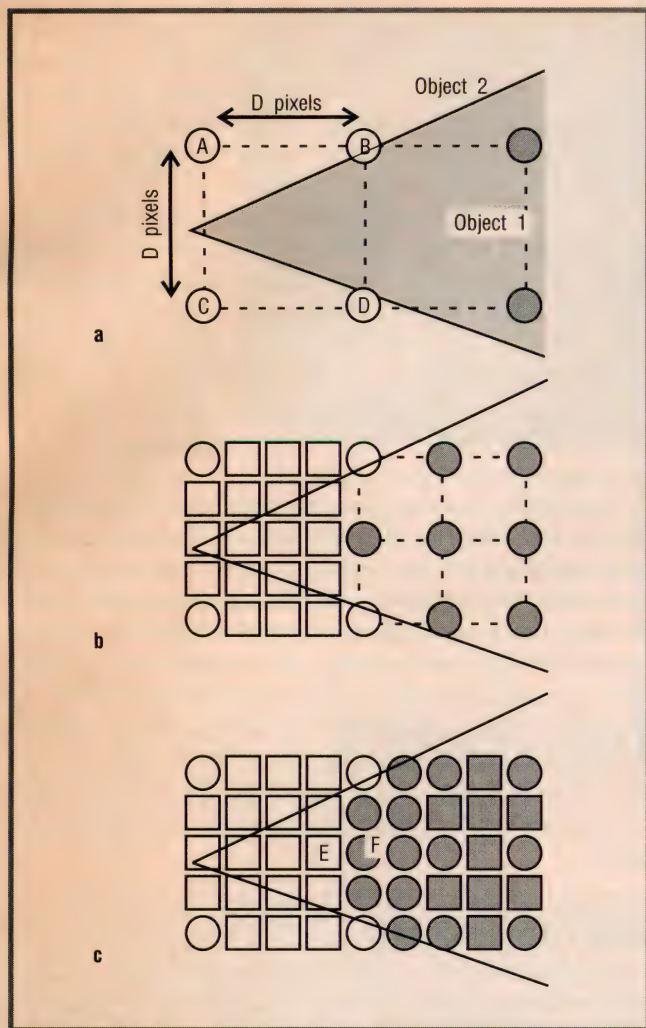


Figure 1. Basic adaptive undersampling algorithm. The open circles indicate pixels calculated by ray tracing; the open squares, pixels calculated by interpolation. Shading indicates the color of Object 1.

undersampling algorithm. PSRT has an improved algorithm that uses pixels with the correct object information from among the sampled pixels to find pixels with erroneous color and correct them. Moreover, PSRT uses ray-object intersection trees for precise classification of the homogeneity of regions and for fast intensity calculation in homogeneous regions.

Image-generation algorithm

Figure 1 shows how a basic adaptive undersampling algorithm works. The algorithm divides nonhomogeneous regions into four small square regions and calculates the intensities of the pixels at each corner of the small square by regular ray tracing, as shown on the right side of Figure 1b.

This basic algorithm is simple and easy to implement. But, as

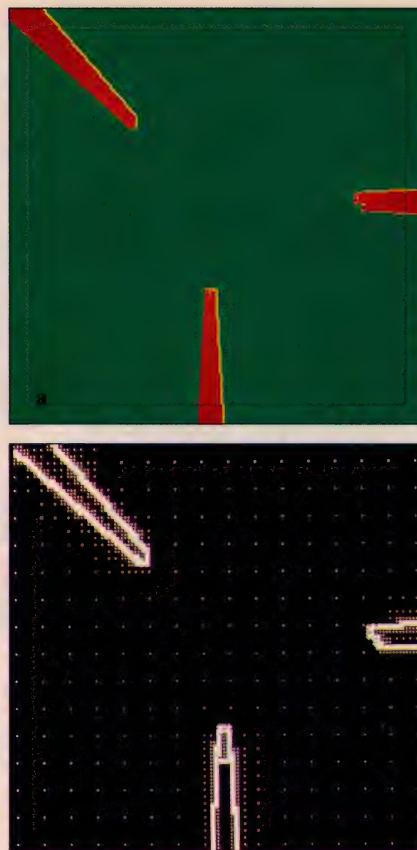


Figure 2. (a) Image generated by the basic algorithm. (b) White dots indicate pixels whose intensities were calculated by regular ray tracing.

Figure 2a shows, sharp corners of objects tend to vanish from the generated images. The white dots in Figure 2b indicate the pixels whose intensities were calculated by regular ray tracing.

Figure 1c illustrates a missing sharp corner. The corner includes a neighboring pixel pair consisting of pixels E and F. Each pixel has a different color, and pixel E is calculated by interpolation. If a pixel pair has two colors, an object boundary should exist between the two pixels in the pair, because the intensity

of one pixel is calculated by interpolation, which could be incorrect. (In Figure 1c, it is incorrect.) The pixel intensity calculated by interpolation must be recalculated by regular ray tracing to ensure the correct intensity.

PSRT detects the conditions illustrated in Figure 1. Using ray tracing at each iteration in hierarchical adaptive undersampling, it recalculates the intensity of pixels incorrectly calculated by interpolation.

Below we describe the image-generation algorithm of PSRT. Figure 3 illustrates how it renders the wedge-shaped object introduced in Figure 1.

PSRT's image-generation algorithm

1. Calculate the intensity of the coarsely sampled pixels by ray tracing, as shown in Figure 3a.
2. For the four pixels located at each corner of the square region (A, B, C, and D in Figure 3b), compare the intensities and the tree structures obtained by ray tracing and classify the similarities into four levels: 0, 1, 2, and 3. (The next section describes the criteria of similarity classification.)

3. Calculate the intensity of the pixel located at the center of the square by regular ray tracing, accelerated ray tracing, or interpolation according to the similarity level determined in Step 2, as shown in Figure 3b. (The next section describes the intensity calculation procedures.) When the center pixel intensity is calculated by regular ray tracing, jump to Step 7 in the error detection and correction procedure.

4. Compare the intensities and tree structures of the four pixels at each corner of the diagonally oriented squares (B, F, D, and E in Figure 3c) and classify the similarities of the region in the same way as in Step 2.

5. Calculate the intensity of the pixel located at the center of the diagonal squares in the same way as in Step 3, as shown in Figure 3c. When regular ray tracing is used, jump to Step 7.

6. Apply Steps 2 through 5 repeatedly, subdividing the region until the intensities of all pixels in the image are determined, as shown in Figures 3f and 3g.

Error detection and correction procedure

7. From the four corners of the region, find the pixel or pixels whose pixel intensity is calculated by interpolation and is dissimilar to the center pixel. Recalculate the pixel intensity by regular ray tracing (E in the region bounded by B, F, D, and E in Figure 3d).

8. If the recalculated intensity is not the same as the intensity before recalculation (E in Figures 3c and 3d), classify the similarities of the corner pixels in the three other regions that contain the recalculated pixel and recalculate the center pixel intensities of these three regions (G, H, and J in Figure 3e).

9. Apply Steps 7 and 8 repeatedly. (In our example, Step 7 is

applied to the regions whose center pixels are G, H, and J in Figure 3d.) If no pixel intensity was recalculated, return to Step 3 or Step 5.

Even if interpolation gives incorrect intensities to some pixels

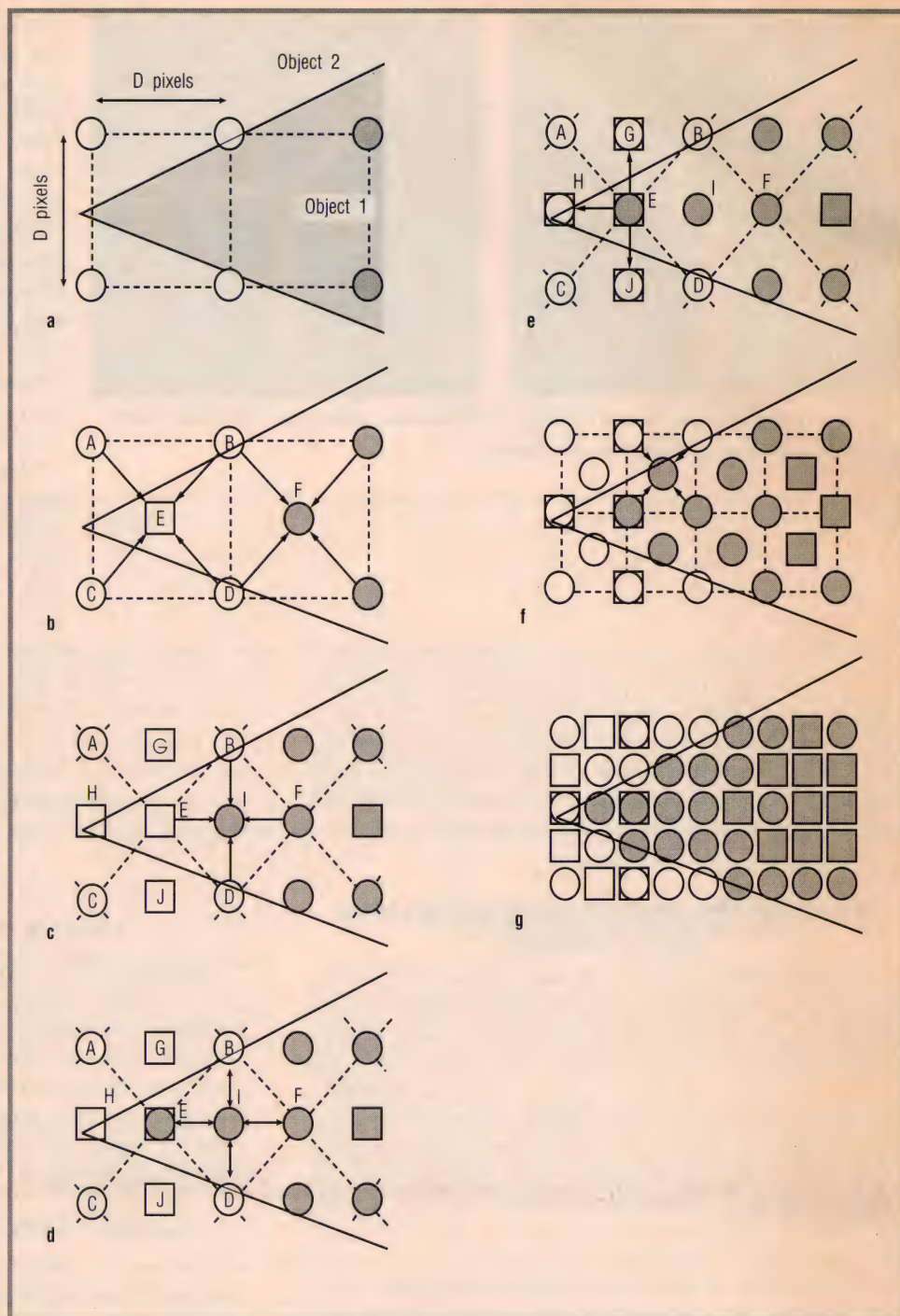


Figure 3. PSRT image-generation algorithm. The open circles indicate pixels calculated by ray tracing; the open squares, pixels calculated by interpolation. A circle inside a square indicates an interpolated pixel recalculated by ray tracing. Shading indicates the color of Object 1.

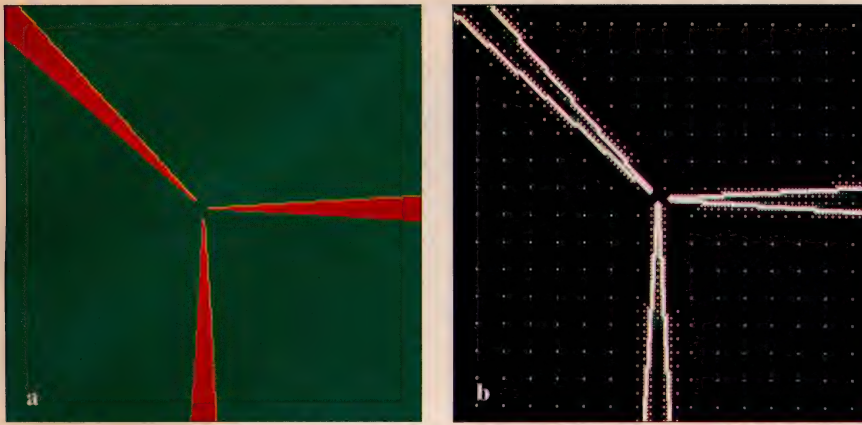


Figure 4. (a) Image generated by PSRT. (b) White dots indicate pixels whose intensities were calculated by regular ray tracing.

in homogeneous regions, execution of the algorithm results in correct recalculation of the intensities, as Figure 3g shows. Figure 4a shows an image generated by PSRT. The white dots in Figure 4b indicate the pixels whose intensities were calculated by regular ray tracing. Comparing Figure 2 with Figure 4 shows how PSRT completely eliminates the problems of parts missing from thin objects or tips missing from wedge-shaped objects.

Similarity check and intensity calculation

We use information from ray tracing to measure the similarities of pixels at region corners. Figure 5 shows transmitting rays and tree structures obtained by tracing the rays. In the figure, P and P' are pixel intensities; $O1$, $O2$, $O3$, and $O4$ are object names; $R1$, $R2$, $R1'$, and $R2'$ are reflected rays; and $T1$, $T2$, $T1'$, and $T2'$ are refracted rays. As the figure shows, we can obtain the pixel intensities and the tree structures of the ray-object intersections. Each node of the tree structure contains the following information:

- The intensity at the ray-object intersection point.
- The identification code of the object intersecting the ray.
- The surface physics of the object (for example, texture, transparency).
- Shadow appearance at the intersection point.
- The intensity of illumination from all light sources at the intersection point.
- The position (x, y, z) of the intersection point.
- The local geometry of the object surface at the intersection point (for example, normal vector, curvature).

By comparing pixel intensities and the tree structures, we can examine the similarities among pixels precisely. If we use the

similarities properly, we can calculate the intensity of the pixel at the region center using the simplest procedure. For example, if the object identification codes in the ray-object intersection tree of the four pixels are the same (that is, the path of ray transmission is the same), we do not need to find the object that intersects with the ray, even if the intensities of the four pixels differ. We can calculate the intensity of the center pixel by ray tracing along the same path. If the intensities of illumination or shadow appearances are the same, shadow calculation is not necessary.

We determine the quantity of information to be recorded and the number of levels of similarities to use on the basis of the available memory and the calculation needed to measure the similarities.

In the experiment that we describe in

next section, we actually use the pixel intensity and the ray-object intersection trees—which contain the object identification code, the attributes of the object, and the shadow appearance—for the similarity check.

The four levels of similarity that we used follow. For each level we give the criteria that determine the degree of similarity among pixels and the simplest pixel-intensity calculation.

Level 0: Active pixels

- **Criterion.** Different objects appear in each tree (paths of ray transmission differ).
- **Intensity calculation with regular ray tracing.** Because there are edges in the region around the four corner pixels and we cannot predict the path of ray transmission through the center pixel, we must calculate the center pixel's intensity by regular ray tracing, including processing to find shadow appearance.

Level 1: Low-moderate pixels

- **Criterion.** The paths of ray transmission are the same, although shadow appearances differ.
- **Intensity calculation with accelerated ray tracing.** Because we cannot predict shadow appearance at the center pixel, we must calculate the precise position of the ray-object intersection point and shadow appearance. However, since we can predict that the ray's path through the center pixel is the same as the path through the corner pixel, we can calculate along the path. A time-consuming search for the object that intersects the ray is not necessary.

Level 2: Mid-moderate pixels

- **Criterion.** The path and shadow appearances are the same. However, objects in the tree have texture attributes,

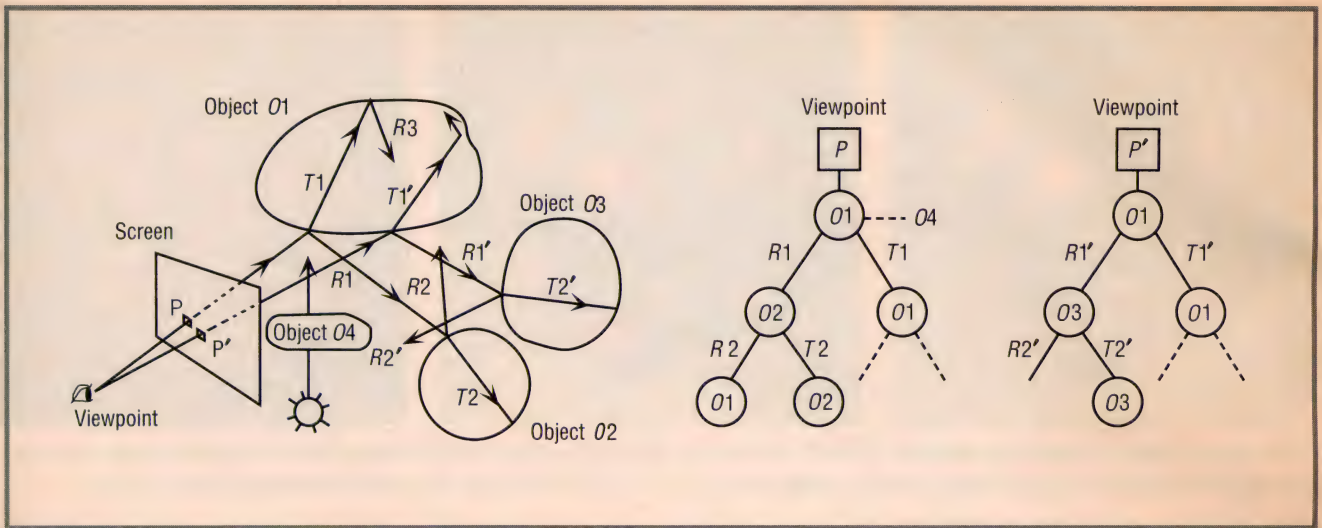


Figure 5. Transmitting rays and their tree structures.

and/or the differences in pixel intensities are large:

$$\max(|\bar{r} - r_1|, \dots, |\bar{r} - r_4|, |\bar{g} - g_1|, \dots, |\bar{g} - g_4|, |\bar{b} - b_1|, \dots, |\bar{b} - b_4|) > T$$

$$\bar{r} = \frac{1}{4} \sum_{i=1}^4 r_i, \bar{g} = \frac{1}{4} \sum_{i=1}^4 g_i, \bar{b} = \frac{1}{4} \sum_{i=1}^4 b_i$$

where $(r_i, g_i, b_i) \{i = 1, \dots, 4\}$ are the intensities of each color component of the pixels, and T is a predefined threshold.

- **Intensity calculation with accelerated ray tracing.** Because we cannot predict changes in pixel intensity in the region, intensity recalculation is necessary. However, we can calculate along the same path and do not need to check shadow appearance.

Level 3: High-moderate pixels

- **Criterion.** The path is the same, the shadow appearances are the same, there is no texture-mapped object, and the differences in pixel intensities are small:

$$\max(|\bar{r} - r_1|, \dots, |\bar{r} - r_4|, |\bar{g} - g_1|, \dots, |\bar{g} - g_4|, |\bar{b} - b_1|, \dots, |\bar{b} - b_4|) < T$$

- **Intensity calculation with interpolation.** Because we can predict that the region contains one smooth object surface, we can calculate the center pixel intensity by intensity interpolation. For example, we calculate the center pixel color by averaging each color component of the four corner pixels.

Reducing intersection calculation

With the similarity procedures of Levels 1, 2, and 3, we can reduce the time-consuming ray-object intersection calculation.

Finding the intersection point between a known ray and a known object becomes the main calculation at Level 2. At Level 1, we need to do this and find any shadowing object against light sources in which shadow appears differ among the four corner pixels.

Experimental results

We generated several images using PSRT and compared processing time and image quality with images generated by standard ray tracing. We wrote the programs in C for a Sun 4/260 and used the bounding spheres acceleration technique in both programs.

Figures 6a, 7a, and 8a show three images generated by standard ray tracing. Figures 6b, 7b, and 8b show the three images generated by PSRT. Figures 6c, 7c, and 8c show the positions of active pixels, low-moderate pixels, and mid-moderate pixels (described in the previous section) as white dots, red dots, and green dots, respectively. Pixel intensities in blank areas were calculated by interpolation.

The size of the images in Figures 6 through 8 is 512×512 pixels. The interval of the initial ray-traced pixels (D) was 4 pixels. The threshold we used to check the differences in pixel color (T) was 16. (Each pixel color component has a value between 0 and 255.)

As the figures show, the PSRT images are visually the same as those generated by standard ray tracing. Figures 6c, 7c, and 8c demonstrate that the most active pixels, except for initial ray-traced pixels, are located near primitives' edges. Low-moderate pixels gather near shadow edges. Mid-moderate pixels are located in the parts where intensities change suddenly or where there are texture-attributed primitives.



Figure 6. (a) Sample image generated by standard ray tracing. (b) Sample image generated by PSRT. (c) White dots indicate positions of active pixels (Level 0); red dots, low-moderate pixels (Level 1); and green dots, mid-moderate pixels (Level 2).



Figure 7. (a) Sample image generated by standard ray tracing. (b) Sample image generated by PSRT. (c) Positions of pixels of Levels 0, 1, and 2.



Figure 8. (a) Sample image generated by standard ray tracing. (b) Sample image generated by PSRT. (c) Positions of pixels of Levels 0, 1, and 2.

Table 1 shows the numbers and types of pixels and image-generation times. Average error is the amount of error per pixel defined by the equation in Figure 9. In the equation, $R(x,y)$, $G(x,y)$, $B(x,y)$ and $r(x,y)$, $g(x,y)$, $b(x,y)$ are the intensities of

each color component of the pixels calculated by standard ray tracing and PSRT, respectively.

As Table 1 shows, we calculated the intensities of many pixels by interpolation and applied time-consuming regular ray trac-

Table 1. Number and type of classified pixels and image generation time by PSRT.

	Figure 6b	Figure 7b	Figure 8b
Number of primitives	plane = 25 sphere = 8 quadric = 12	plane = 6 sphere = 4 quadric = 13	plane = 2 sphere = 7 quadric = 7
Active pixels* (regular ray tracing)	24,046 (9.2%)	30,624 (11.7%)	35,158 (13.4%)
Low-moderate pixels* (accelerated ray tracing)	4,170 (1.6%)	5,275 (2.0%)	15,380 (5.9%)
Mid-moderate pixels* (accelerated ray tracing)	1,923 (0.7%)	74,046 (28.2%)	92,765 (35.4%)
High-moderate pixels* (interpolation)	232,005 (88.5%)	152,199 (58.1%)	118,841 (45.3%)
Recalculated pixels	51	214	936
CPU time (Sun 4/260)	431s	678s	1,211s
Relative speed**	9.0	4.6	2.0
Average error (intensity range= 0~255)	0.668	0.241	0.058
*Total number of pixels per image = 512 x 512 = 262,144			
**Relative speed of PSRT compared with standard ray tracing.			

$$\text{average error} = \frac{\sum |R(x, y) - r(x, y)| + \sum |G(x, y) - g(x, y)| + \sum |B(x, y) - b(x, y)|}{3 \times \text{number of pixels}}$$

Figure 9. Equation defining average error per pixel in Table 1.

ing to only a few pixels. In this experimental environment ($D = 4$, $T = 16$), PSRT was two to nine times faster than standard ray tracing.

Table 2 shows the distribution of the number of pixels with various error values. The error of most pixels is relatively small; only a few have significant error.

In PSRT, the amount of error and image-generation speed vary according to interval D and threshold T . Figure 10 shows the relative speed and average error for various intervals D and thresholds T , when Figure 7 was generated by PSRT.

As Figure 10 shows, when the interval D or the threshold T increases, the average error increases. However, the increase in the relative speed is small compared with the increase in threshold T and is also small when interval D becomes long, such as $D = 8$, 16, or 32. Results were similar for Figures 6 and 8.

These results demonstrate that long interval D and high threshold T are not necessary for fast image generation. We should select relatively short intervals and low thresholds to

avoid missing small objects and producing conspicuous interpolation error. This is why we generated the images in Figures 6 through 8 with $D = 4$ and $T = 16$. Our experience shows that with PSRT the interval $D = 4$ and the thresholds $T = 8$ or $T = 16$ produce images visually nearly the same as images generated by standard ray tracing.

In the experiments reported here, we used 512×512 -pixel images and no anti-aliasing. For generating smaller images—say, 256×256 or 128×128 pixels—with the same primitives, PSRT's ray-tracing acceleration effect would probably decrease, because the percentage of pixels that could be calculated by interpolation or accelerated ray tracing would decrease. For generating smaller images with a very large number of primitives, the speed of PSRT reduces almost to that of standard ray tracing. However, generating smaller images by ray tracing is not a big problem with today's faster computers.

For the generation of larger antialiased images having many pixels—say, $4,096 \times 4,096$ or $8,192 \times 8,192$ pixels—PSRT is very effective. In such cases, we regard 512×512 -pixel images as tiny parts or subpictures of the larger image. We need

Table 2. Distribution of the number of pixels for various errors.

Error	0	1~3	4~7	8~11	12~15	16~19	20~23	24~27	28~31	32~
Figure 6b	75,715	185,772	617	36	4	0	0	0	0	0
Figure 7b	208,875	52,146	922	137	50	11	3	0	0	0
Figure 8b	240,189	21,542	213	119	57	10	7	0	0	7

many primitives to express the larger images, even if the complexity (for example, the total border length of image components) of each subpicture differs little from the examples in Figures 6 through 8.

Conclusions

Because very small objects can vanish from images generated by PSRT, this method may not always be suitable for extremely high quality image rendering. However, the degradation of

image quality in our experiments was very low when we used relatively short intervals of initial ray-traced pixels. Faster image generation results from combining PSRT with a method that reduces ray-object intersection calculation, making PSRT very useful for applications where speed is important. □

Acknowledgments

We thank Yukio Kobayashi, the general director of the Visual Perception Laboratory at NTT Human Interface Laboratories, for his encouragement and overall support. We also thank Akihiko Hashimoto and all the members of the Visual Perception Laboratory for their many helpful suggestions.

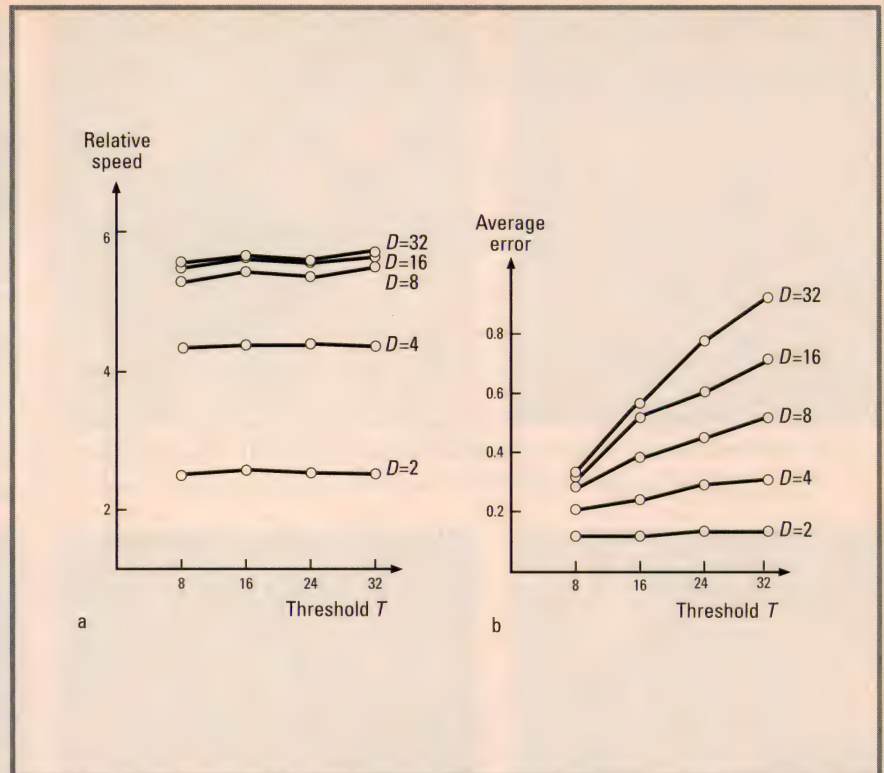


Figure 10. Relative speed (a) and average error (b) for various values of threshold T .

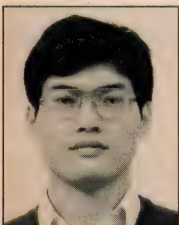
References

1. T. Whitted, "An Improved Model for Shaded Display," *Comm. ACM*, Vol. 23, No. 6, June 1980, pp. 343-349.
2. M. Shinya, T. Takahashi, and S. Naitou, "Principles and Applications of Pencil Tracing," *Computer Graphics* (Proc. Siggraph), Vol. 21, No. 4, July 1987, pp. 45-52.
3. M. Lee, R. Redner, and S. Uselton, "Statistically Optimized Sampling for Distributed Ray Tracing," *Computer Graphics* (Proc. Siggraph), Vol. 19, No. 3, July 1985, pp. 61-67.
4. D. Mitchell, "Generating Antialiased Images at Low Sampling Densities," *Computer Graphics* (Proc. Siggraph), Vol. 21, No. 4, July 1987, pp. 65-72.
5. D. Thomas, A.N. Netravali, and D.S. Fox, "Anti-Aliased Ray Tracing with Covers," *Computer Graphics Forum*, Vol. 8, No. 4, Dec. 1989, pp. 325-336.



Takaaki Akimoto is a research engineer in the Visual Perception Laboratory at Nippon Telegraph and Telephone's Human Interface Laboratories. He joined NTT in 1984. His research interests include algorithms and hardware architecture for image synthesis.

Akimoto received a BS and an MS in information engineering from the Kyushu Institute of Technology, Japan, in 1982 and 1984, respectively.



Kenji Mase is a senior research engineer in the Visual Perception Laboratory at NTT Human Interface Laboratories. He joined NTT in 1981 and was a visiting researcher at the MIT Media Laboratory from 1988 to 1989. His research interests include image processing, computer graphics, computer vision, and their applications in the human-machine interface.

Mase received a BS in electrical engineering and an MS in information engineering from Nagoya University, Japan, in 1979 and 1981, respectively.



Yasuhito Suenaga is a senior research engineer and supervisor in the Visual Perception Laboratory at NTT Human Interface Laboratories. He leads a project researching computer analysis and synthesis of human images. He joined NTT in 1973 and was a visiting researcher at the MIT Media Laboratory from 1985 to 1986.

Suenaga received his BS, MS, and PhD in electrical engineering from Nagoya University, Japan, in 1968, 1970, and 1974, respectively. He is a member of the Institute of Electronics, Information, and Communication Engineers and the Information Processing Society of Japan.

The authors can be reached at NTT Human Interface Laboratories, Nippon Telegraph and Telephone Corp., 1-2356 Take, Yokosuka, Kanagawa, 238-03 Japan. Akimoto's e-mail address is akimoto%nttcvg.ntt.jp@relay.cs.net

THE FOLLOWING INFORMATION IS AVAILABLE:

Contact the Publications Office; to facilitate handling, please request by number.

- Membership application, student #203, others #202
- Periodicals subscription form for individuals #200
- Periodicals subscription form for organizations #199
- Publications catalog #201
- Compmail electronic mail brochure #194
- Technical committee list/application #197
- Chapters lists, start-up procedures #193
- Student scholarship information #192
- Volunteer leaders/staff directory #196
- IEEE senior member grade application #204 (requires ten years practice and significant performance in five of those ten)

To check membership status or report a change of address, call the IEEE toll-free number, 1-800-678-4333. Direct all other Computer Society related questions to the Publications Office.

PURPOSE

The IEEE Computer Society advances the theory and practice of computer science and engineering, promotes the exchange of technical information among 100,000 members worldwide, and provides a wide range of services to members and nonmembers.

MEMBERSHIP

Members receive the acclaimed monthly magazine *Computer*, discounts, and opportunities to serve (all activities are led by volunteer members). Membership is open to all IEEE members, affiliate society members, and others interested in the computer field.

1951-1991

40 YEARS OF SERVICE



IEEE COMPUTER SOCIETY

A member society of the
Institute of Electrical and Electronics Engineers, Inc.

PUBLICATIONS AND ACTIVITIES

Computer. An authoritative, easy-to-read magazine containing tutorial and in-depth articles on topics across the computer field, plus news, conferences, calendar, interviews, and product reviews.

Periodicals. The society publishes six magazines and five research transactions. Refer to membership application or request information as noted above.

Conference Proceedings, Tutorial

Texts, Standard Documents. The Computer Society Press publishes more than 100 titles every year.

Standards Working Groups. Over 100 of these groups produce IEEE standards used throughout the industrial world.

Technical Committees. More than 30 TCs publish newsletters, provide interaction with peers in specialty areas, and directly influence standards, conferences, and education.

Conferences/Education. The society holds about 100 conferences each year and sponsors many educational activities, including computing science accreditation.

Chapters. Regular and student chapters worldwide provide the opportunity to interact with colleagues, hear technical experts, and serve the local professional community.

OMBUDSMAN

Members experiencing problems — magazine delivery, membership status, or unresolved complaints — may write to the ombudsman at the Publications Office.

EXECUTIVE COMMITTEE

President: Duncan H. Lawrie*
University of Illinois
Dept. of Computer Science
1304 W. Springfield
Urbana, IL 61801
(217) 333-3373

President-Elect: Bruce D. Shriver*
Past President: Helen M. Wood*

VP, Standards: Paul L. Borrell (1st VP)*
VP, Press Activities: Barry W. Johnson (2nd VP)*
VP, Conferences and Tutorials: Laurel V. Kaleda†
VP, Education: Raymond E. Miller†
VP, Membership Activities: Ronald D. Williams†
VP, Publications: Ronald G. Hoelzeman†
VP, Technical Activities: Mario R. Barbacci*

Secretary: James H. Aylor*
Treasurer: Joseph Boykin†
Division V Director: Edward A. Parrish, Jr.†
Division VIII Director: Helen M. Wood*
Executive Director: T. Michael Elliott†

*voting member of the Board of Governors
†nonvoting member of the Board of Governors

BOARD OF GOVERNORS

Term Expiring 1991:

P. Bruce Berra, Michael Evangelist,
Ted Lewis, Raymond E. Miller, Earl E. Swartzlander, Jr.,
Joseph E. Urban, Thomas W. Williams

Term Expiring 1992:

James H. Aylor, Alicja I. Ellis, Tadao Ichikawa,
C.V. Ramamoorthy, Sallie V. Sheppard,
Harold Stone, Akihiko Yamada

Term Expiring 1993:

Fiorenza Albert-Howard, Jon T. Butler, Michael C. Mulder,
Yale N. Patt, Anneliese von Mayrhauser,
Benjamin W. Wah, Ronald Waxman

Next Board Meeting

November 1, 1991, 8:30 a.m.
Opryland Hotel, Nashville, TN

SENIOR STAFF

Executive Director: T. Michael Elliott
Publisher: H. True Seaborn
Director, Conferences and Tutorials: Anne Marie Kelly
Director, Finance and Administration: Tod S. Heisler
Director, Board and Administrative Services: Violet S. Doan

COMPUTER SOCIETY OFFICES

Headquarters Office

1730 Massachusetts Ave. NW
Washington, DC 20036-1903
Phone (202) 371-0101
Fax: (202) 728-9614

Publications Office

10662 Los Vaqueros Cir.
PO Box 3014
Los Alamitos, CA 90720-1264
Membership and General Information:
(714) 821-8380
Publication Orders: (800) 272-6657
Fax: (714) 821-4010

European Office

13, Ave. de L'Aquilon
B-1200 Brussels, Belgium
Phone: 32 (2) 770-21-98
Fax: 32 (2) 770-85-05

Asian Office

Ooshima Building
2-19-1 Minami-Aoyama, Minato-ku
Tokyo 107, Japan
Phone: 81 (3) 3408-3118
Fax: 81 (3) 3408-3553



IEEE OFFICERS

President: Eric E. Sumner
President Elect: Merrill W. Buckley, Jr.
Past President: Carleton A. Bayless
Secretary: Hugh Rudnick
Treasurer: Theodore W. Hissey, Jr.

VP, Educational Activities: Richard S. Nichols
VP, Professional Activities: Michael J. Whitelaw
VP, Publication Activities: J. Thomas Cain
VP, Regional Activities: Robert T. H. Alden
VP, Technical Activities: Fernando Aldana



Guest Editors' Introduction: Multimedia—It's Actually Useful!

Jack Grimes
Consultant

Mike Potel
Apple Computer

Coming up with a good definition for multimedia engaged us in quite a struggle. The sheer number of possibilities we encountered prompted us to write "What Is Multimedia?" for the January 1991 10-year retrospective issue of *IEEE Computer Graphics and Applications*, put together last fall. We had already decided that multimedia was "useful stuff," committed ourselves to this special issue, and corralled several people who agreed to write articles. That article could have served as a lengthy introduction for this issue. But, even as the ink dried on the January article, we had new thoughts on multimedia. Is this a fast-moving field or what?

Then, in April 1991, the *Communications of the ACM* issue on digital multimedia systems hit our mailboxes. That magazine makes a great companion issue for the one in your hands. Whereas this *CG&A* issue focuses on applications, the *CACM* issue covers three different areas: digital image and video standards, hardware for video encoding and decoding, and future standards. Edward Fox's introduction also includes background sources and a glossary to help you interpret the alphabet soup. We think both special issues make great reading. The articles in this issue of *CG&A* provide specific examples of multimedia's usefulness. However, a key question remains: Will multimedia become mainstream?

Will multimedia become mainstream?

It depends. A fundamental problem afflicts multimedia authoring—not enough people have the necessary skills. Whether this even constitutes a limitation or not depends on your model of how people will use the more elaborate forms of multimedia, those incorporating sound and motion video.

We believe two main models cover the possibilities pretty well:

- User-centered: The producer is also the consumer is also the producer.
- Traditional: There are few producers and many consumers.

The user-centered model, with the producer as consumer, draws parallels with the use of personal computers in general and desktop publishing in particular. The enormous power of the tools means that previous consumers of newsletters, flyers, or typeset pages suddenly become producers of their own typeset pages. The user-centered model predicts that consumers of sound and motion video will take up their recorders and camcorders and produce music and home videos. Although not necessarily of the highest professional quality, these productions will be good enough for most purposes, will become very popular, and will certainly provide a new outlet for the creative talents of many current personal computer users.

The traditional model separates producer and consumer. It assumes that the producer needs special, uncommon creative and production skills. This model argues that powerful tools are necessary but not sufficient to make acceptable multimedia productions. This model draws analogies with the record industry and the movie industry, where specialized people exercise their craft and produce CDs and movies for the rest of us to consume—and consume we do. Further, it argues that we all have built-in standards for acceptable quality, based on CDs, broadcast television, and Spielberg's films. Given this expectation for quality, we will not voluntarily show our home videos publicly (unless we hope to win a \$10,000 prize).

If the user-centered model holds true, multimedia will certainly become mainstream. Likewise, the traditional model might also spawn a very large multimedia industry.

Depending on which model proves correct, multimedia will touch us as individuals in different ways. Will we (user-centered) use multimedia products in business, like other outside-supplied technology such as the laser printer? Or will we (traditional) all purchase CD-ROM drives for our computers, buy titles, and dramatically improve the world's level of education by our own efforts? Prospectively, both models involve very large use of the technology. It's the near-term prospects that vary between them.

Under the user-centered model, progress will likely be slow. It takes five to 10 years for a new technology to become adopted to the extent that it is pervasive. For example, consider graphi-

cal user interface (GUI) technology. It got its major start with the Smalltalk work at Xerox PARC in the late 1970s and became commercially viable in 1984 with the Macintosh. Here we are in 1991—and the marvel of the last 12 months was the strong showing of GUI on the PC and compatibles! Even more telling, the majority of personal computers still use a text-based interface. (You probably wouldn't have known that, since you are obviously an enlightened graphics person to be reading this magazine. . . .)

Under the traditional model, multimedia could become an overnight success, like the Nintendo phenomenon. Since the creative talents only need to reside in a few people, and since the traditional distribution channels can obviously handle rapid adoption, progress could be astounding. If it doesn't take off, on the other hand, then it might remain a niche capability reserved for keynote addresses at Comdex and Macworld.

Which way will it go? Stay tuned, you have found the right channel. □

Acknowledgments

This issue is filled to overflowing with articles on applications of multimedia. This means that the *CG&A* editors, the authors, and especially Beth Fox at Apple, worked very hard to coordinate and complete this special issue. Thank you.



Jack Grimes is currently a consultant and author. Previously, he was Vice President of Technology and Vice President of Engineering at MASS Microsystems, where he was responsible for the selection and acquisition of new technology by the company and for product development. Before that, he was with Intel as a product line architect responsible for graphics chips.

Grimes received his BS, MS, and PhD in electrical engineering from Iowa State University in 1965, 1968, and 1970, respectively, and an MS in experimental psychology from the University of Oregon in 1981. He is a member of IEEE, ACM, and the Human Factors Society. He has served on the editorial boards of several IEEE and ACM journals and is the past technical editor of *Computer*.



Mike Potel is Director of Software Engineering at Apple Computer. His group is responsible for future system software and development tools for Macintosh personal computers.

Potel received his BS in mathematics from the University of Michigan in 1970 and his MS and PhD in information sciences from the University of Chicago in 1971 and 1977, respectively. He is a member of IEEE and ACM and is on the editorial board of *IEEE Computer Graphics and Applications*.

Readers may contact Grimes at 780 Galli Dr., San Jose, CA 95129, e-mail jgrimes@cup.portal.com or Potel at Apple Computer, 20525 Mariani Ave., MS:37/MP, Cupertino, CA 95014, e-mail potel@apple.com or Apple Link Potel.



Multimedia Scouting

Kristina Hooper Woolsey
Apple Computer Multimedia Lab

These Multimedia Lab projects show how uninformed viewers can get interactive access to elements formerly reserved for graphics, film, and music professionals.

The debut of Hypercard in 1987 made a range of sights and sounds available to the typical viewer. In addition to providing a general linking tool and a computer language (Hypertalk) nonprogrammers could comprehend, Hypercard also provided easy access to optical media—particularly videodiscs and CD-ROMs. These optical media gave the viewer quick, inexpensive access to visual and aural elements from the worlds of television, movies, and music. But the big question was “What will the average viewer do with these media-rich, interconnected materials?” Apple established its Multimedia Lab in 1987 to address this question—to do the initial “scouting” in this new territory.

Over the last three years, the Multimedia Lab has experimented with a range of techniques and developed several examples, often in collaboration with other organizations. The context for these experiments was developing new media types through the serious consideration of existing media. Citing some of those examples here, I examine six existing media types and describe how we extended them into new genres of information display.

Developing new media

Many years ago, Doug Engelbart^{1,2} asserted that computers could augment human thinking and communication and therefore assist people in dealing with the increasing complexities of their worlds. Multimedia's specific area of interest is the incorporation of interactive elements—pictures, movies, diagrams, and sounds—into a learning environment that supports exploration, presentation, and construction activities.

To achieve this goal, the Multimedia Lab has built on existing landmark media and experiences. For example, we asked "What happens when we add movies to text?" "What if we incorporate text and audio interviews with movies?" "What does a multimedia activity look like?"

The platform we used to explore these possibilities typically consisted of a Macintosh computer attached to a CD-ROM and a videodisc player. In some cases we used Macintosh SE computers and two screens, one for computer display and the other for video. In other instances we used Macintosh II computers and, with the use of video display boards, single screen systems that displayed both computer and video information. These configurations were not exotic; they were off-the-shelf items. Yet the experiences they delivered were profound, helping us to predict future user scenarios with multimedia displays. They have contributed to our new platform developments, which incorporate sound and video more conveniently and inexpensively, making these techniques more accessible to customers—and thus, ubiquitous in our culture.

Gathering existing media

In trying to design computer environments that augment human activity, we considered how people gather information, both in and out of classrooms, in electronic and nonelectronic environments, in public spaces as well as in private surrounds.

We identified six basic classes of activities. The first of these involve print materials (including atlases, historical documents, encyclopedias, novels). Even in the most automated office or classroom, these print materials are highly visible and well-used.

The second category is the use of audiovisual displays. This category includes slide shows, documentaries, movies, and television. These are now widely available to computer users through the incorporation of optical media. Although generally used for entertainment, these displays are also employed frequently and effectively in classrooms and homes.

Speeches constitute the third category of presentation. This oral tradition continues to permeate our information gathering habits, even though it is "old fashioned." Business meetings and classroom presentations are typical of this genre.

Activities present a more interactive environment for viewers. Games, simulations, puzzles, and design tasks are all examples of activities. In these instances, we learn not by simply listening or collecting, but by direct personal experience with a task.

Another arena that emphasizes active participation is the materials tradition of the kindergarten classroom, a tradition



Figure 1. In *Worldview*, a mouse click on the map reveals a movie of a flyby over the highlighted area.

that infuses many teaching approaches. It's based on the premise that there is a general set of "stuff" that viewers must explore and manipulate to understand a topic.

Self-expression is our final category. Consistent with a constructivist tradition, we assert that individuals learn a great deal in the act of expressing themselves. Currently most people do this in text or spoken words. We are interested in how self-expression is extended when nonprofessionals gain access to images and sounds to create their own compositions. Recently we've become enthused about making such self-expression conversational in multimedia.

These six categories provide a framework for explaining the multimedia examples we built between 1987 and 1990. Though arbitrary at some level, they establish some initial bearings from which to explore totally new media types.

Print and multimedia

What happens when you add video or film imagery to printed materials? Can maps be enhanced with images and sounds? Can imagery bring the text alive for the viewer? Does music enhance the investigation of a theme? Might news clips provide a context for understanding a novel?

The answer to these questions is "yes." We developed a number of examples in which we enhanced the effectiveness of existing text by adding multimedia elements.

One of the first Hypercard design examples was *Worldview*, a prototype of an electronic atlas (shown in Figure 1). To use it, the viewer clicked on a map to see a picture of the environment represented. The system brought the "territories" represented by the map alive. Such a display not only encouraged interest in a particular place but also helped people to understand how maps work. It took a series of squiggles on a page and related them to elements in the real world. A cross on the map became

the Eiffel Tower or the Golden Gate Bridge. A line became a flyby over the Nile.

The reference aspects of the atlas were still available to the viewer. Viewers could look up the population of Paris while viewing the French countryside or explore the food industries of Africa while looking at a desert landscape. The pictures provided a context for the interpretation of data, and the computer gave the viewer quick access to the particular topic. So what had been just a mechanism to gather pictures from a map became a more complex information environment, one easily guided by the user's interests.

It was relatively simple to build such a resource with existing technology, and users found it highly engaging. One teacher said that it was what she had always hoped computer education would be. (See the "Project Credits" sidebar for more information on this and the other projects mentioned in this article.)

A more pervasive print resource is the encyclopedia. We worked with Groliers Publishing to bring this print resource alive with multimedia (see Figure 2). In this case we chose a set of encyclopedia articles that addressed western expansion in the United States. To make the articles more engaging, we added stories about individuals traveling west. These stories guided the viewer through the materials with a particular point of view.

In addition, the design team developed a set of overviews accompanied by music. These oriented the viewer to existing

materials and introduced them to new ones. The quickness of access to the articles and the multimedia integration helped individuals find answers to their questions and provided a highly textured information environment in which they could explore ideas easily.

Another challenge we accepted was to make the ideas of the U.S. Constitution come alive. This text document is important to understand, yet it is frequently inaccessible to the general user or student. In our *Constitution* project we incorporated movies and diagrams to extend the original text document. Simple time line organizations emphasized the idea that the Constitution is a living document, changing as our culture changes (see Figure 3). We showed how different amendments were added to the document over the years. Viewers could access these changes by choosing specific dates or topics. Students and teachers alike found this presentation highly involving and a good framework in which to raise critical constitutional issues.

This project helped us realize how important general tools would be in multimedia compositions. We found, for example, that viewers wanted to develop their own pathways through materials. Some wanted to view the materials, mark places of interest, then follow their trail repeatedly. In the context of this project, we developed a "trailmaker" tool. We began to see that access to print materials was just the first step. Viewers wanted tools to assist them in gathering and presenting their own materials.

Project Credits and Information

For *Worldview* (1987) Fabrice Florin was the lead designer.

For the *Groliers Encyclopedia* project (1988), the lead designers were Tim Oren, Kristee Kreitman Rosendahl, Gitta Soloman, and Peter Cook.

For *Constitution* (1987), a collaboration with Scholastic Publishing and Optical Data, the lead designers were Fabrice Florin and Steve Gano.

Pat Hanlon, Bob Campbell, and Steve Gano were the lead designers for *Grapevine* (1987).

Life Story (1988), a collaboration with the Smithsonian Institution, Lucasfilm, and Adrian Malone Productions, was designed by Fabrice Florin, Rob Semper, and Ed Bastian. Adrian Malone, Steve Arnold, and I were the executive producers.

Interactive Nova: Animal Pathfinders (1988) was a collaboration with WGBH and Peace River Films. The lead designers were Bill Purdy and Ted Sicker, and Barry Cronin, John Borden, and I were the executive producers. The product is now available through Scholastic Publishing in New York.

Designed by David Lawrence, *Disappearing Ducks* (1988) was a collaboration with Lucasfilm and the Audubon Society.

Geography Television (1986) was a collaboration

with the National Geographic Society and Lucasfilm. The prototype designers were David Lawrence, Doug Crockford, and Matthew Kerns. *GTV* is now available through Optical Data Publishing, Warren, New Jersey.

Multimedia Presentation Resource (1989) was produced by Burt Arnowitz, with thanks to Sueann Ambron, Steve Gano, and Kristee Kreitman Rosendahl.

The Visual Almanac (1989), designed by the Apple Multimedia Lab, is now a noncommercial product available to educators through Optical Data Publishing.

School Restructuring (1989) was a collaboration with the American Federation of Teachers. Lead designers were Marsha Levine, Bruce Goldberg, and Bill Purdy.

Future Worlds (1989), a collaboration with the BBC Interactive TV Unit, was designed by Steve Gano, Kristee Kreitman Rosendahl, and Max Whitby.

Fabrice Florin was *Moss Landing's* lead designer (1989).

In 1990 Margo Nanny and Bob Mohl led the *Aspen Kids* project. The designers were Charles Kerns, Steve Gano, Kristee Kreitman Rosendahl, Nancy Hechinger, and Fabrice Florin.

Unless otherwise noted, I was the executive producer/director for these projects.



Figure 2. A wide range of points of view can complement existing text resources, as shown in this view of *Grolier's Encyclopedia*.

My last print-based example is *Grapevine*.^{3,4} This project centered around John Steinbeck's novel, *The Grapes of Wrath*, and involved Pat Hanlon, a high school English teacher, and Bob Campbell, the school's librarian. The issue was that of context. Students unfamiliar with 1930s America, the Depression, and agriculture found *The Grapes Of Wrath* difficult to identify with. Yet when Pat and Bob provided students with news clips from the Depression era, photographs of dust bowl farms, and the music of the period, the novel came to life. Individual students related it to their own lives and their experiences of dislocation from their homes, economic hardships, and labor issues. The multimedia "surround" to this novel enhanced its ability to speak across the years.

An impressive aspect of the *Grapevine* project is that it was assembled by educators, not professional developers. Hanlon and Campbell mastered Hypercard programming to articulate their ideas in their own environment, rather than relying on the interpretation of technological professionals. This has encouraged other educators to create similar projects for themselves and their associates.

From atlases to novels

In all the examples I have described, the primary form—an existing print document—remained intact. Multimedia elements just elaborated on it. Yet as people experience these projects, they find that the initial print materials become less prominent, and the multimedia elements begin to dominate. Viewers remember the rousing civil rights speeches of James Meredith, the shape of the Eiffel Tower, or the faces of migrant workers. And they recall their interaction with these images as they selected some of them for their own presentations. Despite our original goal of simple enhancement, the result is a complex, powerful experience for the viewer.

Now imagine products *designed* to combine print materials with other media elements. At the very least, this might create

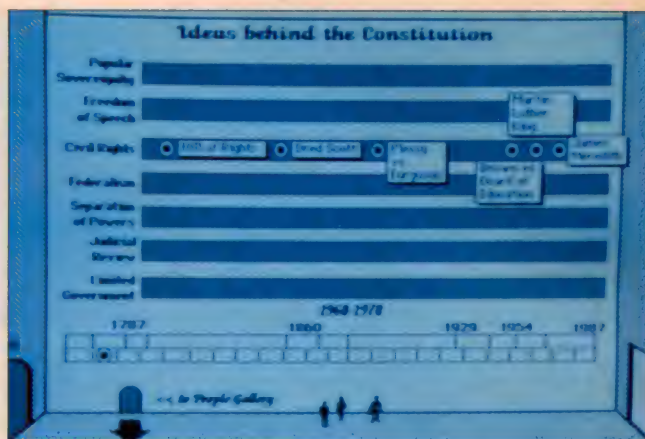


Figure 3. This theme-oriented time line gave viewers access to a range of documents relevant to the U.S. Constitution.

a more effective communication of ideas. At the most, it could produce exciting new media types. If the success of the projects I have described here is any indication, similar products will soon find their way into the marketplace.

A number of encyclopedia manufacturers already produce electronic versions based on existing text, with enhanced search capabilities and minimal multimedia elements. Textbook developers are experimenting with ways to enhance their text-oriented displays. And innovative teachers are creating Hypercard-based multimedia presentations that enliven and enrich their curricula.

Audiovisuals and multimedia

Audiovisual materials challenge us to design a computer environment cordial to video. A range of design issues arises from this task. For example, how can the emotional elements of movies and documentaries work effectively within the neutral, logic-based environment of the computer? How can we inform by using techniques developed for entertainment? Can a combination of music and still images engage viewers, rather than remind them of their experiences with educational slide shows?

Life Story is one of the archetypes in this category. We took one of BBC's dramatic movies, *Life Story: The Discovery of DNA*, and wove ancillary materials around it. For example, we added documentary interviews with the actual scientists portrayed in the movie. We provided text transcripts of scenes so that people could review the materials, including a glossary of some scientific concepts covered in the film. We created simple simulations of some concepts. And we referenced other materials—articles, books, movies—of interest for viewers who wanted to know more. Finally, we produced a navigational template for all these diverse elements, based on a "pamphlet" structure. This allowed viewers to delve into the wealth of materials rather than drown in them.

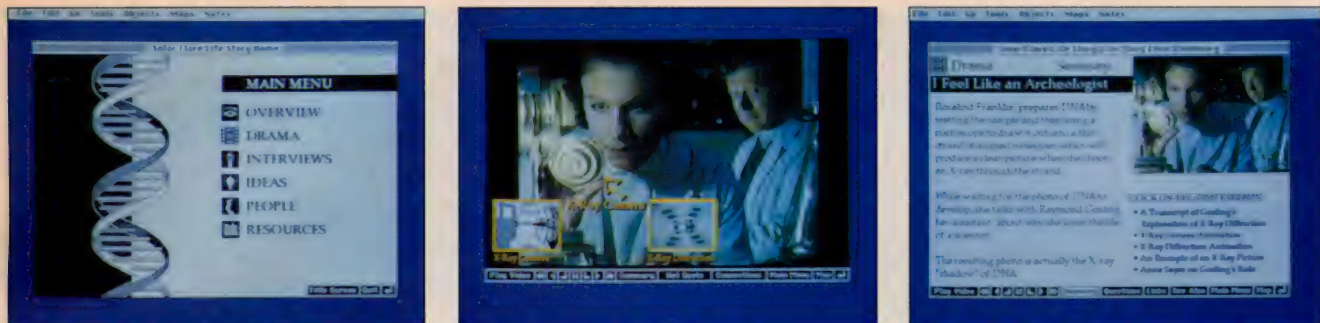


Figure 4. (a) The *Life Story* prototype gave viewers access to the study of DNA from a number of perspectives, including scenes from a drama and interviews with scientists. (b) Viewers who chose to enter from the drama were provided with elaborations about individual scenes. (c) Additional resources about people and objects in key scenes could be gathered by clicking on them.

We expected that people would first view the movie *Life Story* from beginning to end in a linear fashion. Then viewers could explore the additional materials in a nonlinear manner, motivated solely by their individual interests (see Figure 4). They might access video interviews with the real James Watson, for example, or run a simulation to learn more about the x-ray diffraction camera shown in the movie. A text description of a particular scene would alert them to other relevant materials—text, movie, sound, or diagram—available in the system. We also provided ideas on how this particular program might be used in the classroom, including movie clips of classroom use and interviews with teachers who used these materials.

Many students who didn't find biology interesting changed their minds after the dramatic impact of this movie. The students who appreciated biology were further stimulated by the

portrayal of the scientists' lives and what it means to "do" science, rather than just study it.

Two other things about this project are noteworthy. First, the pamphlet structure proved very successful in helping people navigate through a large multimedia database. It served as a general model for other projects, particularly those designed around movies. Second, allowing viewers to click directly on a video frame to access related information proved to be a compelling experience. Controlling movies by pointing directly at their elements gave people a sense of personal power.

For the *Interactive Nova: Animal Pathfinders* project, we took a nature film about animal migration and added materials that would turn it into a substantial part of a biology curriculum. As in *Life Story*, students would view the original movie first, learning about the migration of different animals. Then, using our materials, students and teachers could broaden their investigations beyond migration to include additional species, different environments, and other animal behaviors (see Figure 5).

A primary use for these resources is teacher and student presentations. During the course of a school year, a teacher can use this multimedia product continually, to discuss different animals, environments, and behaviors. Substantial materials were developed for this purpose, including extensive text documentation. In addition, we included simple tools that allowed students and teachers to edit video and text elements for their own presentations.

The product also contained three video-centered activities. One taught students how to interpret bee dances, testing their comprehension of basic directions via film clips of bees "dancing." A second let students emulate actual investigations of monarch butterflies. We intended to show students what it takes to "do science," rather than passively learning from the work of others. A final activity taught students about the potential extinction of turtles by participating in a detective adventure. Here the "spirit of the chase" gave context to short video segments describing the perils turtles face today.

Many multimedia projects are basically "click and see." Viewers choose which video segment they want to see and in

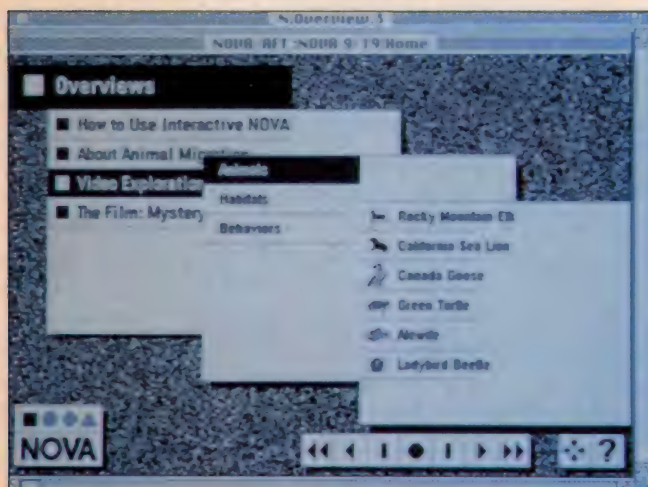


Figure 5. *Interactive Nova: Animal Pathfinders* is now a classic model for a multimedia product, containing a range of content materials, activities, and tools. In the Overview, viewers can choose to see the complete movie, short introductions to the material, or instructions.

what order. With good video segments this can be a compelling environment for learning, especially if the segments are short. The construction of *Interactive Nova* elaborated on this basic model and included diverse activities as well as tools for construction.

We chose a different model for our *Mystery of the Disappearing Ducks* project (shown in Figure 6). We acquired raw video footage developed for a TV special on wetlands. We wanted to take advantage of materials on the cutting room floor. Accordingly, we paid attention to the ideas behind the movie (and how we could apply them), but not to the final form of the movie itself. This project presented the challenge of tailoring other people's footage to our purposes. It proved a complex and idiosyncratic enterprise.

However, it also provided some opportunities. For example, we were able to include more in-depth materials than the television show. We presented controversial topics from a range of perspectives. By comparing and contrasting arguments among farmers, conservationists, and hunters, ecology came alive.

The *Disappearing Ducks* example was designed by a class of high school students on the premise that they would know how to develop something engaging for their peers—and indeed they did. Working with professional designers, the students created an entire mystery around a befuddled character named Paul Parkranger. A naturalist, Paul had left his cabin with an unsolved case of vanishing wetlands on his desk. It was the students' task to take over the case and solve it.

The student design team learned a great deal about ecology in the process of creating this game. (It provides a good basic model of how students or teachers might create a multimedia product to learn a particular topic.) The mystery itself proved engaging and informative. Though based on a particular nature documentary, our program encouraged exploration far beyond the original materials. There was also no trace of the original documentary's linearity.

My final example in the audiovisual category is *GTV* or *Geography Television*. This commercial product resulted from a collaboration between Apple Computer, Lucasfilm, and the National Geographic Society. Our task was to combine imagery and computers to create a geography program "hip enough" to engage the imaginations of middle school students.

The initial design of this project focused on still images for a number of reasons: the number of historical images available as stills, the National Geographic's experience with slide shows, and the technological possibility of delivering this product on CD-ROM. More intuitive reasons included a desire for "MTV-like" images (a cross between animation and live-action video) and an assumption that movement could be depicted through music rather than the actual images.

We also naively thought that still imagery would be easy to manipulate. As it turned out, we had to develop a customized video editor for our Macintosh II platform. Most video editing systems are tailored for editing motion sequences, not prolonged stills. A crisp relationship between images and sound was required for the music to "drive" the stills. (The final prod-

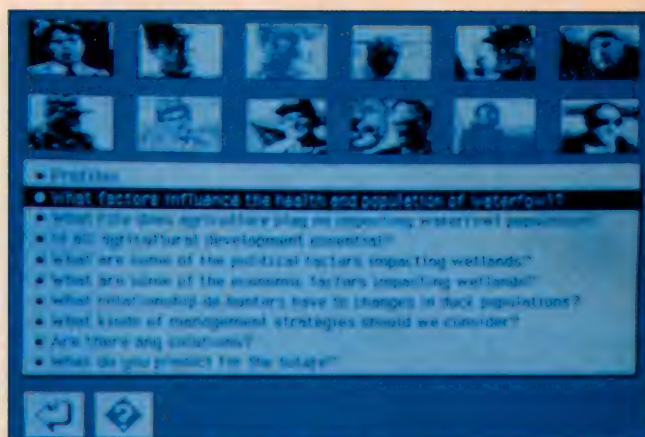


Figure 6. One section of the *Disappearing Duck* project enabled students to witness the environmental arguments set forth by different interest groups. They clicked on a question, then a person, to hear the perspectives of these people.

uct includes some movie sequences as well.)

Activities and tools for creating presentations added to *GTV's* appeal, allowing viewer experience to extend beyond the image sequences and materials provided. This new type of "slide show" has been successful in classrooms. Short sequences give teachers great "conversation starters" that can be integrated with other classroom materials. And they have none of the stodgy feeling many of us associate with the hum of a slide projector.

Integrating analog and digital

The marriage of video and computer technologies is a natural one. People enjoy manipulating live-action sequences and are easily engaged in explorations that move them from video to text to diagrams and so forth. New by-products of this marriage should emerge quickly. The trick will be how to finance an industry whose elements and technologies are new and fluid. Engineers must resolve the conflict between the basic analog nature of video signals and the digital nature of computer displays. Going "all digital" means grappling with the huge storage and memory requirements of simple analog video segments and developing a range of compression schemes for current use. Yet the effort being directed toward these technical challenges indicates that the compelling combination of these signal types will soon result in a wide range of support.

Speeches and multimedia

Personal computers are typically thought of as one-person machines, used for word processing, business analysis, or games. Recent observations have proven that they can also be collaborative centers for small group interactions. The emergence of Hypercard and its control over media introduced an entirely new idea: the personal computer as a resource for large group presentations.

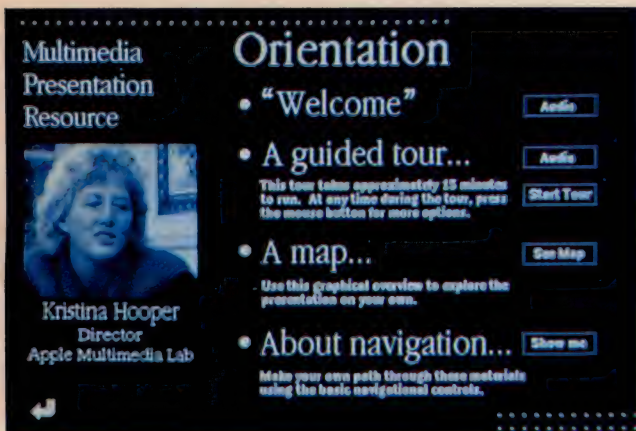


Figure 7. The “Multimedia Presentation Resource” provides speakers with an orientation to the interlinked text and video materials available. Speakers can thus use this system to prepare their own multimedia presentations.

During Hypercard’s introduction at Mac World in 1987, we prepared a speech on the use of Hypercard with multimedia. Instead of clicking on slides or turning on videotapes, we developed a Hypercard program that let us display some of the multimedia examples we had produced. Inadvertently, we created a performance environment, by illustrating how a speech could be enhanced with media elements. We didn’t substitute media for a speaker, the way a flashy multi-image slide show or videotape does. Instead, we gave the speaker an added dimension, which improved the effectiveness of a person-centered presentation.

We quickly realized the implications of such an environment. For example, we needed a remote control to deliver the presentation, as many stages lacked the space to use a mouse as a pointing device. (I’ve given too many speeches with my back to the audience and my mouse careening across a podium.) Another problem was technical support; speeches are typically delivered in a variety of cities and conference centers. Hired or in-house tech crews are usually knowledgeable about slide and video shows and sometimes even large-screen projection. Yet even today, there are very few crews who can assist in setting up a computer-centered multimedia presentation.

Then there was the issue of flexibility. We needed multimedia presentations that could be produced by a variety of nonmedia professionals and could also be modified easily for different audiences. This is important if multimedia is to become common in board rooms and conference centers.

We field-tested some of these principles with a presentation titled “Multimedia Presentation Resource.” This example began innocently enough in spring 1989 as a custom presentation created for Sueann Ambron (cofounder of the Multimedia Lab) and I to brief the press on Multimedia. The original version offered an orientation on different aspects of multimedia, including some history, examples of current work, and implica-

tions of future developments. We then took on the challenge of transforming our customized version into a flexible, public presentation, suitable for a variety of audiences and presenters (see Figure 7).

It took us months to achieve the conversion. We had to develop a navigational structure for the presentation, something that allowed multiplicity of direction in a consistent manner. We needed to make the software more robust and eliminate some bugs. More critically, we needed to provide novice presenters with some background on the materials in the presentation—something the original speakers had not needed, because they had personally developed them. Finally, although people were to modify the presentation as they saw fit, we included our ideas for an effective delivery, encouraging speakers to build from our experiences rather than from scratch.

The final presentation was much simpler than the original. It consisted basically of two parts, both accessible to the presenter in a menu. We added a front end, describing just what the presentation was and how it might be used. We included an audio “guide function,” which demonstrated one possible presentation format and commented on some of the materials.

In addition we created a “behind the scenes” element to complement the basic materials. This included text descriptions of the materials and listings of video elements, so they could be directly accessed from any “location.”

The “Multimedia Presentation Resource” has been used by many Apple employees over the last year. The format has been successful, but the currency of information is a drawback. As a culture we demand that speeches—in contrast to other media—be both highly personalized and up-to-the-moment, especially in technology. While some of this presentation’s materials are “classics,” others have quickly become outdated. Some crucial recent developments are not even included. It will be interesting to see whether other topics—such as education—that do not demand such up-to-date information will prove more suitable for predeveloped presentation formats.

In any case, computer-centered presentations are gaining momentum. They are currently used in classrooms, business meetings, and large-scale conferences. A number of different remotes have been developed to streamline the speaker’s task. Tools and templates allow “normal professionals” to create group presentations from scratch. Or you can employ a media group to assist you in producing a presentation. Some conference centers even have computer presentation setups, so you can arrive with just your data and leave the hardware at home.

Custom presentations, public resources

Speeches are a challenge to multimedia designers. Their task is to enrich the presentation with background materials, yet not allow these materials to eclipse the speaker. You should be able to use your presentation from the podium and modify it on the spot in response to different audiences or even a particular question. Different speakers should be able to use the material, to brief themselves on a topic beforehand as well as to present it to others.

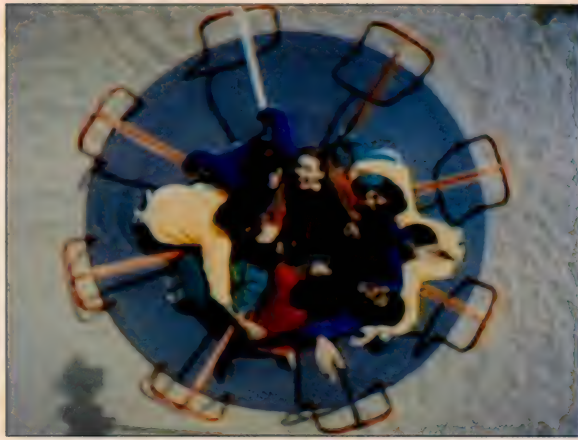


Figure 8. In this merry-go-round activity, viewers watch kids moving in and out on a playground merry-go-round. On a separate screen, a graph charts how the speed of the merry-go-round changes with this movement.

Tools are another part of the challenge. They should be simple, allowing quick, easy assemblage of presentation material. Currently, this process is far from trivial, unless one happens to have access (in themselves or on their staff) to programming, graphic design, and video-editing expertise. However, the situation is improving; tools are being produced that allow users to acquire minimal skills in a relatively short period of time.

The present ratio of custom multimedia presentations to public multimedia presentations is about 100 to 1. The ratio between presentations created by "media professionals" and "nonmedia professionals" is currently about 25 to 1. If we reverse these relationships, we will have a new communication form in our everyday culture; if we don't, we will gain another professional media delivery system—something that is exciting, but not culture transforming.

Activities and multimedia

Computer-generated activities include "what-if" scenarios, simulations, and a plethora of video adventure games. Computer educators have recognized all of these for their learning potential. They are cited as ways to develop new classes of thought through the interaction of visual images and abstract concepts.

In our analyses of multimedia activities, we paid attention to these computer activities. We also looked more generally at other activities, for example the games we played as children. The general idea was to examine those moments where you do something, which is responded to, which drives other activities, and so forth. In other words, we examined situations when opportunities are generated afresh, based on past actions and some random variables. We found a number of opportunities to involve multimedia technologies and materials within this activities framework. Some of the most compelling examples are included in our *Visual Almanac* project.

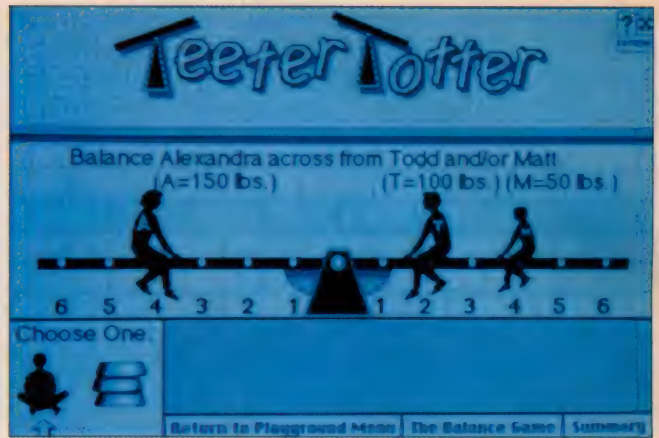


Figure 9. In the Teeter-Totter activity, viewers control video with the computer to determine how to balance a teeter-totter with people of different weights. (Another view of this activity appears at the top of page 26.)

We developed the *Visual Almanac* as an "Interactive Multimedia Kit," illustrating how current technologies could be used to experience and create multimedia examples. It contains 7,000 multimedia objects (still images, movies, and sounds) and some compositional tools, so users can select, retrieve, and manipulate these objects into their own presentations. The product also includes 14 "activities," designed explicitly to show the wide variation possible in multimedia activities.

Among them is "Playground Physics," a set of three activities illustrating how video segments can enhance learning about physical principles. It draws on viewers' past experiences in playgrounds to provide a relevant context.

One activity within "Playground Physics" is the Merry-go-round sequence. The principle it conveys is the conservation of angular momentum. In this activity, a group of children are on a spinning merry-go-round. They move in and out on the merry-go-round, making it go faster and slower. Just watching this is informative. However, viewers are simultaneously presented with a graph of the ongoing activity, displaying how the speed of the merry-go-round changes over time. They can click on the graph (or on an accompanying data record) to see where the kids are at any moment. The answer to the question "where are the kids when the merry-go-round is going the fastest?" for example, is a picture of all the kids clustered together tightly at the center of the merry-go-round, accessed by clicking the highest point on the graph (see Figure 8).

Students and teachers enjoy this question and answer process with pictures, especially their anticipation of the answer while viewing the different events. They also enjoy repeating the sequence, comparing the answers to their own intuitions. Many assert that the pictures evoke their own memories of being on a merry-go-round. Thus they are convinced that this is a "real" phenomenon.

Another "Playground Physics" activity we developed is the Teeter-Totter sequence (see Figure 9). This has the spirit of a

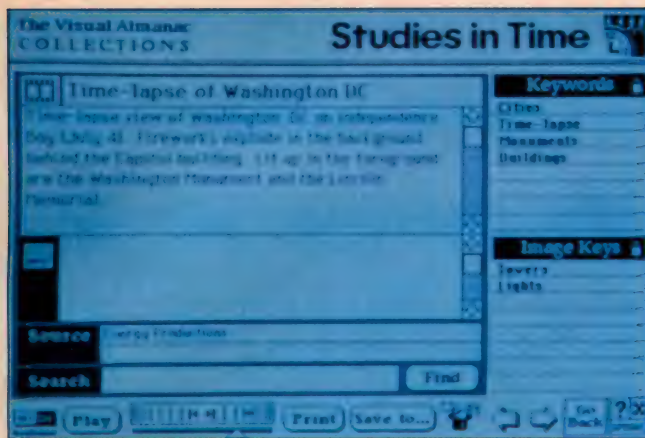


Figure 10. The Visual Almanac provides viewers with access to collections of images, a composition workspace in which to manipulate them, and some sample activities. Each element in a collection is a “multimedia object.” Data cards like this one display the attributes of each object.

simulation, as viewers can change the position of three people (of different weights) on a teeter-totter to achieve balance. As they move people around, the video display shows how the seesaw will tilt (or balance).

Most viewers have a general feel for this problem and move the people around to all sorts of locations. They know there must be a rule for balancing these people and try to discover it through trial and error. Many are astonished to find that a single equation relates all possible combinations/locations of these weights. Just as surprising is the discovery that a useful equation can be fun.

A very different activity in the *Visual Almanac* is “Durations.” Here we used original and acquired footage of time-lapses and made them into activities by giving the viewer playback control of the imagery—backward, forward, and rate of speed.

The time lapses we chose are varied. They include rapid events that can be slowed down, such as a bullet hitting an egg or a milk drop. There are also slow events that can be sped up, such as a group of starfish feeding, or the sun setting, or the evolution of humans.

While many “Durations” seem silly, like fruit rotting or an ice cream sundae melting, it is still compelling to control these events, running them forward and backward (does fruit really emerge from rotten goo?) and at different speeds. Teachers have found a range of uses for these activities, from creative writing subjects to arithmetic calculations to discussions of “seeing the unseen.” Giving viewers interactive control of visual records is apparently both engaging and informative.

Other activities in the *Visual Almanac* include a “Historical Atlas” that chronicles U.S. migration patterns, an orchestra where you can select sounds of different musical instruments, and a “Planetary Highway” that lets you drive through the solar

system, experiencing the spatial relationships between planets. All these encourage viewers to explore, either through the selection of alternatives, the development of theories, or the manipulation of materials.

Materials and multimedia

The “materials tradition” is well established in schools, particularly at kindergarten and elementary levels. It centers around the concrete nature of materials and their importance in developing abstract thought. The Montessori schools, among others, are deeply embedded in this tradition. Libraries are a more abstract model of this tradition, providing teachers and students with “stuff”—images or texts or audiotapes—that can be used in different ways.

The materials tradition drove the basic structure of the *Visual Almanac*. We felt that computer users should have access to a diverse set of images and sounds to make their own multimedia compositions. We wanted to create a present-day simulation of a future probability: personal, computerized access to massive libraries of sights and sounds.

The basic unit for the almanac is the *collection*. Each collection contains hundreds of multimedia objects, each consisting of an image and any relevant data: its title, description, source, and keywords (see Figure 10). Focusing on “the world around us,” we developed 14 collections, including American History, Animals and Plants, Studies in Time, and Everyday Life.

Viewers can search for these images in a number of ways. They can choose a collection and browse its contents. Or they can select a keyword and retrieve the images that match it. Or they can type a word, and find all the objects including that word in their title or description. They can also choose a “rabbit in a hat” icon and jump randomly from one object to another.

Once chosen, an image can be used in different ways. Users save it to a selections list for later viewing or use the composition tools to drop it into a new presentation. (More about that in the next section.)

Teachers and students find this form of image database useful and informative. Sometimes they browse through the images randomly for curiosity and pleasure. Or they search collections for specific ideas, such as “What can I use to inspire a class discussion about John F. Kennedy?” or “Is there anything here to show Mars?” One topic frequently sparks an idea for another. For example, browsing images of Japan can lead to images containing Japanese children and from there to children of other countries. The *Visual Almanac*’s images are highly varied and capable of being used in a wide range of contexts. In this sense, they are fine examples of the materials tradition.

A second project in this same tradition was *School Restructuring*, produced in collaboration with the American Federation of Teachers (AFT). Our task was to give teachers, administrators, and parents information about restructuring: changing the basic structures of America’s education system. Over the years, the AFT had gathered videotapes and articles on this topic. They were interested in making these materials available to a

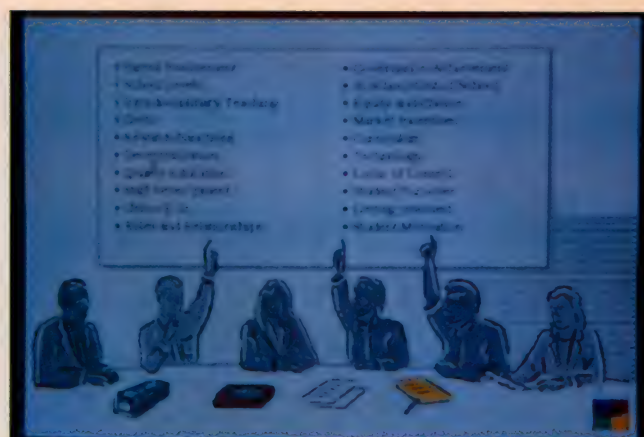


Figure 11. (a) The four sections in the restructuring project provides access to different views of this theme; (b) In the Issues section, when you click on an issue, individuals around the table raise their hands. Viewers select which person's perspective they want to hear. Relevant resources are available through the icons at the bottom of the screen.

wide constituency to encourage informed involvement in restructuring across the country.

The model for this project was a dialogue, because restructuring is an ongoing activity and therefore not amenable to singular analysis, recommendations, and solutions. We wanted to give people background materials so they would understand the issues and participate effectively in restructuring discussions in their local areas.

The materials are grouped into four basic sections (shown in Figure 11a). The first is the Topics section, giving people the basic "gist" of restructuring. As they select particular topics—"Why Restructure?" for example, or "Roles and Relationships"—they hear a short audio overview about the topic. Connections between a particular topic and other sections in the system are also indicated, so that they can jump to any of them for further elaboration.

The second section is Case Studies. Here viewers learn what is going on in schools across the country. As viewers choose a particular school's location, relevant materials from any of the four sections are brought to their attention.

A third section is Issues. This section encourages group discussions among viewers. We recorded a small number of people setting forth their opinions on restructuring issues—the role of parents and administrators, for example, or the assessment of students. When viewers select an issue, the people at the table raise their hands (see Figure 11b). The viewers can select one of them and hear their opinion, then follow up with a viewers' discussion. (As in all other sections, they could access additional articles or interviews in the system that elaborate on an issue.)

Finally, in the Stories section, viewers can choose from a set of people, each of whom has a story to tell. This incorporation of storytelling's oral tradition proves helpful in enticing viewers to recount their own experiences.

The pilot for this project has just begun, but it is already

valuable to people involved in restructuring efforts. It engages them in conversations and arguments beyond the materials provided in the system.

Imageware and general resources

As viewers learn to access multimedia materials and manipulate them with tools, the demand for multimedia databases will increase. The key will be to develop standard formats, so users can go across different collections and categories to gather material. The multimedia-object format of the *Visual Almanac* worked quite well in this regard.

Image retrieval will also have to be routine. In the *Visual Almanac*, we relied on text descriptions of images for retrieval. Though this was effective, it was also time-consuming, as most images are not currently attached to text. An alternative might be image-processing techniques that categorize images through visual similarities. This ability is particularly valuable when searching large numbers of images.

In the AFT project we realized that a straight materials tradition (for example, "here is some stuff") is not always appropriate. Though people enjoyed using a library that included images as well as texts, they often had no context in which to approach or evaluate these materials. By making four categories of approach to the same basic materials, we gave viewers a way to appreciate and connect all the materials.

Expression and multimedia

The ability to gather information, form ideas, and present them in an imaginative, coherent fashion is one of the most effective ways to enhance learning. Simple word processing in schools has demonstrated that the quality and quantity of student writing increases with the use of computers. Much of this can be attributed to the easy revision of materials afforded by computers. The challenge now is to extend this success into the

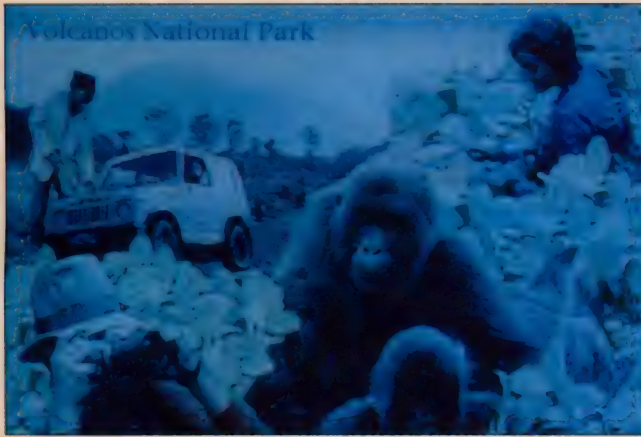


Figure 12. In *Future Worlds*, viewers are presented with sets of multimedia objects relevant to ecological issues. Here, for example, cut-out pictures of tourists, farmers, animals, and scientists in the Volcano National Park represent different topics or themes. The man in the jeep acts as a guide, introducing these topics and how they are interrelated. (Also on page 26.)

domain of multimedia composition, allowing students and others to use media elements as well as text in showing their ideas.

This is a significant task. When enhancing text and graphics, you can still rely on established production conventions and even assume a certain literacy concerning the medium—especially text. (Though the “keyboard issue” often comes between an idea and its expression, this can be overcome.)

In contrast, few people have skills in video and sound production. It is not included in basic curricula, and use of these media is generally confined to professionals, typically on high budgets, for entertainment products. Few of us have experience in using multimedia to casually express our ideas, much less teach these ideas to others. The *Visual Almanac* is our attempt to provide this experience, through the development of simple composition tools and workspaces.

To create a multimedia composition, users browse the *Almanac*’s collections and select a series of images. They then choose one of a number of “card” templates. To this they add text and any of their selected images. Software tools allow users to edit the images (for example, movies can be shortened or run backwards). They can also add sound: bits of narration, music, or sound effects. Hypercard also supplies a paint program, so users can integrate graphics into the composition.

We had a number of nonmedia professionals make compositions with these tools, and some of them are included in the *Visual Almanac* as examples for other users. In one case, a nine-year old boy made a composition about animals. Although similar in structure to a paper and pencil school composition, it has the added dimension of images that enliven the text. Another person created “Presidents,” a crossword puzzle where the clues for presidents’ names are made of images. Other compositions include a trip to space, a fanciful restaurant, California missions, and even the Cuban missile crisis.

Since the *Visual Almanac*’s release, users have elaborated on our examples to create a wide range of imaginative compositions. They enjoy selecting and manipulating images to express their ideas.

Another of our projects oriented to individual compositions is *Future Worlds*, created in collaboration with the Interactive Television Unit at BBC. For this project we gathered a series of interviews with different “future” experts. We then integrated this material into a multimedia system. From this, viewers could create their own vision of the future.

To accomplish this, we had to consider how the experts’ materials would appear in a multimedia environment. We chose to treat each expert’s set of materials as a space composed of multimedia objects, much like the *Visual Almanac*. These collections of objects—typically shown in a single screen—were perused by viewers, then incorporated into individual compositions about the future (see Figure 12).

An important element was the viewer’s ability to create video records from the video elements provided in the experts’ spaces. To accomplish this, a general video editing tool was created (a tool that served as the prototype of Macromind’s Media Maker tool). This tool provided visually based, direct manipulation of video for naive users. A simple interface allowed viewers to rearrange the sequencing of segments, thus creating their own linear record.

This project—like the *Visual Almanac*—made it possible for viewers to rearrange supplied materials to present their own ideas. However, what would happen if they created their own visual records rather than using those provided by others?

This question was addressed in a number of projects, including *Moss Landing*. Here we asked a variety of professionals to record their perspectives on a small California coastal town. They included television newspeople, scientists, and a range of photographers from different disciplines: nature, underwater, surrogate travel, and cinema verité documentary. Our goal was to see what each of these perspectives contributed to the recording of this town. From there, we could decide which approach to suggest to novices, and what to put together to browse these materials.

One insight gained through this project was that filmic materials are time-consuming to log and organize—activities that are important if you want to make sensible selections. Though we spent only three days in the field, it took months to organize what we had gathered and give it coherence.

We ended up with a linked database of video materials for viewer exploration. This included a “hyperpicture” that gave access to different parts of the town from a single viewpoint. Viewers could identify and link the different segments as they chose. Keywords enabled them to base their exploration of segments on a particular theme or element.

Most of all, experiencing the town through an interview with a local girl, a walk through the fish market, the sounds of a salt marsh, or the view from a helicopter was highly compelling. Even if you had never been to the town, you started to understand its “reality” and feel a part of it (see Figure 13).



Figure 13. (a) We compiled a range of materials about Moss Landing, a small town in California. These included video segments describing different locations. **(b)** Linked to each location was a set of other materials.

Moss Landing was a major video production involving professionals. Our next step was using this experience in a project where novices recorded themselves and their environment for their own perusal and comment.

The result was our *Aspen Kids* project, a workshop at the Annual Aspen Design Conference in 1990. About twenty children, 10 to 18 years old, participated in this workshop. They were there to learn about multimedia design. We were there to determine if children could use cameras to gather images and data in the field and then manipulate them as multimedia objects in a computer environment, all within three days.

On Day 1, the children created multifaceted representations of themselves: each one consisted of a short movie, an audio file, a set of five still images (which were shot to imply some animation), and some information (such as “How much do you like homework?” or “Where do you live?”). As in the *Visual Almanac*, all the data for each child was organized on a data card. The children designed their own cards, furthering the individualization process (see Figure 14a).

On Day 2, the kids extended this exercise to the gathering of multimedia objects from the world. Their task was to walk around the town of Aspen, find letters of the alphabet composed of everyday objects (for example a telephone pole is a “T”, a small pond is an “O”), and record them. They made a short movie about each object, using a still image, some relevant sounds, and information about the object.

On Day 3, they viewed the results of their activities in a computer environment. (The Multimedia Lab staff input all the materials in the evenings.) On a simple level they could flip through the data cards, viewing information they found interesting. Or they could take advantage of the computer’s processing capabilities and search for images via keywords, for example, selecting all of the objects that matched these keywords (see Figure 14b).

At a more complex level they could use some of the display

software designed for this workshop. One of these was a Grapher. With the Grapher, viewers could look at people or objects based on different attributes. For example, they might choose the relationship between the height and weight of all workshop participants, or their preferences for classical music versus rock music. After each query, pictures of each participant would align themselves on a graph according to their attributes. Though many questions were silly, the kids learned a great deal about the nature of graphing, and they enjoyed making way-out hypotheses about particular relationships.

Another software function was a Typer. This let them take the letters they had gathered in the field and put them together into sentences. Needless to say, the notion of spelling words with objects was very enjoyable. Also, since each letter was actually a multimedia object, viewers frequently became interested in the letter itself, researching further information about it.

Aspen Kids was a design task, conducted to explore how quickly an imaginative set of materials could be compiled, and whether kids would find such a task involving. Simple changes can make it an “educational” investigation. One model for this is the field trip, where students gather objects that represent different biological species in a terrain, for example, or different art objects in a museum. Or it can be a representation of one’s own environment—family members, local stores, and streets. Cameras and other data-gathering devices bring personal objects into the computer world to explore, manipulate, and compose with. These can complement the more public objects acquired from professional, outside sources.

Conclusions

Wide use of multimedia elements in our culture has impressive implications. We can learn entirely new ways to express ourselves, ways in which images and sound fit into daily discourse. We can understand the uses of media in communication



Figure 14. (a) In the *Aspen Kids* project, kids represented themselves as multimedia objects in a range of media forms, all accessible from their data card. (b) They could view these objects according to values on various attributes attached to each multimedia object (for example, a person) in the Grapher.

and better interpret those media messages (for example, on television news broadcasts) around us. We might even discover how to bring new ideas into public discourse, ideas not amenable to other forms of media.

As the availability of multimedia composing increases, there will be great demand for engaging images and sounds. This demand will come from our inability to be everywhere at once, as well as our inability to capture professional images and sounds. Multimedia might be the "clip art" of the future, allowing us to integrate professionally produced audiovisual elements into our forms of communication at low cost.

Even more interesting is the prospect of viewers' using consumer equipment—such as camcorders—to gather their own materials. The development of general templates and standard data formats would allow people to structure this material into highly personalized compositions. Opportunities for enhanced learning and communication can be boundless.

I have only reflected a sampling of the multimedia experiments we have conducted in the Multimedia Lab in recent years. Our conclusion from these experiences is that people find well-designed multimedia presentations engaging, imaginative, and useful, both as viewers and participants.

Multimedia is clearly important in defining new classes of published works. It enhances the existing traditions of print and film by extending their scope, thus generating a media form somewhere between the two. It adds a dimension of flexibility and pertinence to speeches and suggests new forms of performance and interaction for group presentations.

Multimedia also promises a transformation of our everyday world, through the enrichment of educational and entertainment activities. By providing a wealth of new materials and the tools to structure them, we hope to see casual, everyday use of multimedia materials. These will result in opportunities to display individual perspectives, heighten participation in local issues, and achieve better understanding of faraway places.

There are indeed a number of hurdles for multimedia to overcome if it is to become ubiquitous. There are technological issues to be addressed as well as substantive economic issues. Yet our experiments have shown again and again that there are compelling human experiences to be delivered and created in multimedia surrounds, and we look forward to sustaining momentum to make these widely available. Our scouting has shown wondrous lands ahead. □

Acknowledgments

Thanks to Cindy Rink for extensive editing help on this document. I also owe thanks to several people who helped with specific projects: Walter Koetke, Jim Lengel, and Beverly Hunter for *Constitution*, the BBC for *Life Story*, Craig Southard and Karla Kelly for *Disappearing Ducks*, Burt Arnowitz for *School Restructuring*, and Peter Armstrong for *Future Worlds*.

References

1. D.C. Engelbart, "A Conceptual Framework for the Augmentation of Man's Intellect," in *Vistas in Information Handling*, Howerton and Weeks, eds., Spartan Books, Washington D.C., 1963, pp. 1-29.
2. D.C. Engelbart with K. Hooper, "The Augmentation System Framework," in *Interactive Multimedia*, S. Ambron and K. Hooper, eds., Microsoft Press, Redmond, Wash., 1988, pp. 15-32.
3. R. Campbell and P. Hanlon, "Grapevine," in *Interactive Multimedia*, S. Ambron and K. Hooper, eds., Microsoft Press, Redmond, Wash., 1988, pp. 157-177.
4. R. Campbell and P. Hanlon, "Hypercard: A New Deal in the Classroom," in *Learning with Interactive Multimedia*, S. Ambron and K. Hooper, eds., Microsoft Press, Redmond, Wash., 1990, pp. 257-286.



Kristina Hooper Woolsey is the director of the Apple Computer Multimedia Lab. She has taught at the University of California at Santa Cruz and the Massachusetts Institute of Technology, and she has also worked as the director of research and development at the Atari Sunnyvale Research Lab. She has extended her training as a cognitive scientist into the areas of picture recognition, imagery, and visual technology.

Woolsey received her PhD in cognitive science from the University of California at San Diego and her AB in psychology from Stanford University.

Woolsey can be reached at the Multimedia Lab, Apple Computer, 3220 Sacramento St., San Francisco, CA 94115.



Numbers— A Medium That Counts

Glorianna Davenport
MIT Media Laboratory

Jonathan D. Harber
Di Va Corporation

As computational carriers of numbers, spreadsheets bring hypermedia to the office. A hypermedia spreadsheet we developed helped us construct a template for the first hypermedia version of business case study.

As the computer industry waits for multimedia consumer demand to build, industry experts look for a successor to Lotus 1-2-3, the application that generated exponential growth in the personal computer industry in the 1980s. We propose that the spreadsheet might also be the vehicle that carries hypermedia to the business desktop in the 1990s.

Numbers: An important medium

Multimedia is generally defined as a user-directed medium that incorporates a range of data types, including text, graphics, sound, audio, video, animation, and other computer programs. Ironically, numbers—basic to computing and widely used in the personal computer since its genesis—seem to have been ignored in definitions of multimedia environments. This omission results in an impoverished information space.

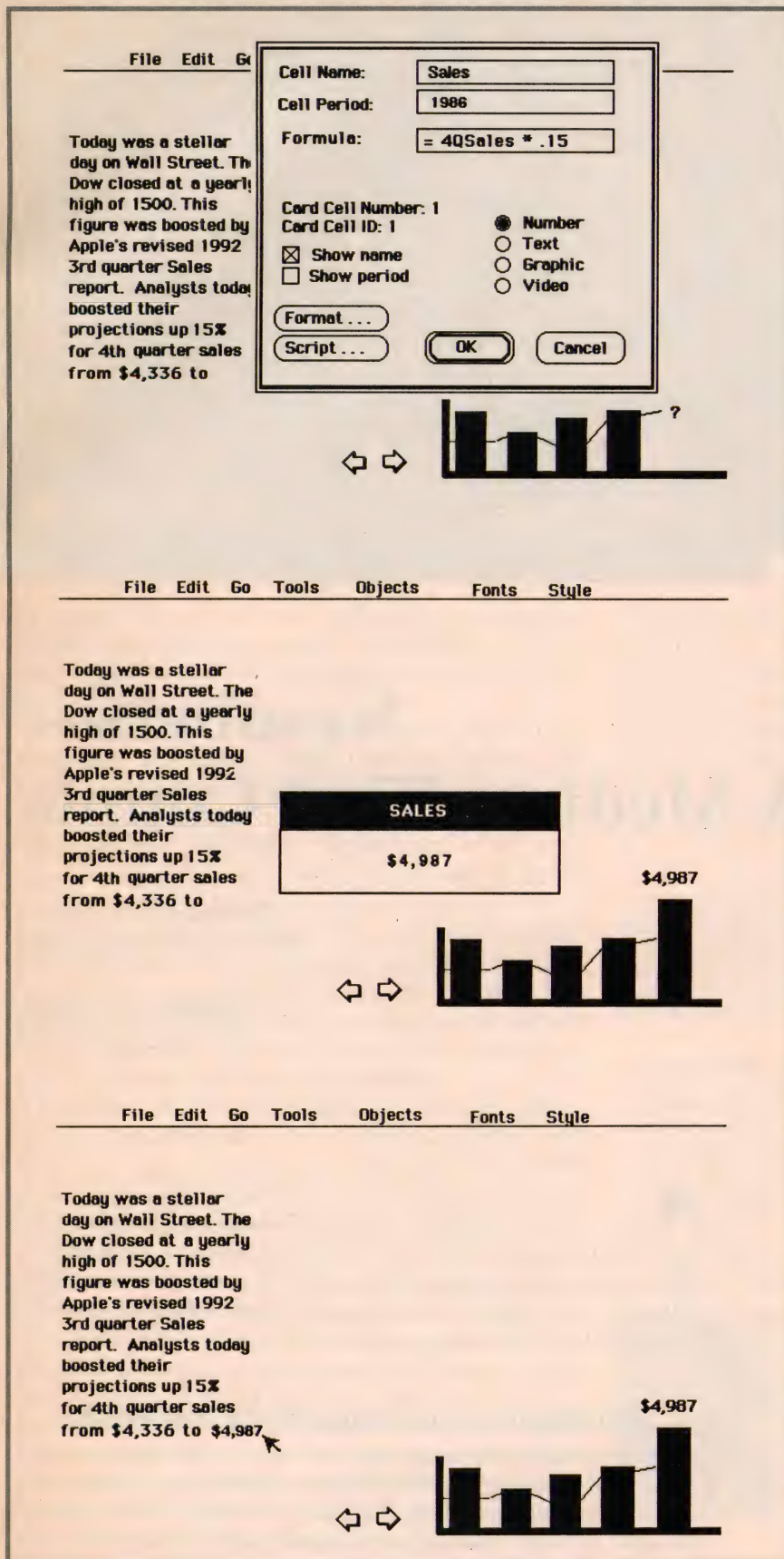


Figure 1. A potential interface for incorporating a cell into a document.

Consider the difference between numbers in a word processor and in a spreadsheet. Numbers are dynamic and have scientific properties that you can use in a spreadsheet. But when you copy a number to a word processor, you lose those properties.

The cell

In 1987 Apple Computer began distributing Hypercard, a software program for Apple's personal computers that gives users the power of hypermedia authoring. Hypercard provides text fields, buttons, and graphics as generic object types. Since 1987 other software companies have incorporated hypermedia capabilities into their products. Unfortunately, most of these products omit the cell.

We propose that the numeric cell is a powerful generic object type. A cell might be considered the smallest unit of computation. As such it has the potential of carrying link information. When creating a document or application, you should be able to create a cell in any location and relate it to other cells or events by using a full set of computational rules or formulas. Figure 1 illustrates a potential interface for incorporating a cell into a document architecture such as Hypercard.

Although hypermedia and multimedia have become somewhat synonymous in recent publications, note that "hypertext," the origin of hypermedia, was first proposed by Vannevar Bush in 1945 as a description of a computerized library system with the ability to search by associations. Reviewing current work in the industry, we might ask "What happened to the hyper in hypermedia?" Associations and interactivity within documents of multiple media require computing capability and some form of programmability. The generic cell could provide this needed link.

If you question the impact of including the cell as a generic object type, consider again the difference between numbers as they appear in a word processor and numbers in a spreadsheet. If a number were truly treated as another data type, it

should have the same properties across applications.

Authoring with numbers

An installed base of sophisticated authors knows how to work in the medium of numbers: the spreadsheet users of the world. Analysts who work with spreadsheets know how to manipulate numbers and operands in a spreadsheet to express subjective analysis of data sets. They understand how to define rules by defining formulas that drive numerical analyses (for example, the Lotus 1-2-3 formula @IF (a2 > a1, 1980, 1990)). As analysts pass spreadsheet files among themselves, the spreadsheet becomes a publication medium. It provides an application to think with; numbers and the computational formulas constitute equal parts of the content carried.

Cells as active agents

Vast quantities of information are becoming available in digital format. As computers become linked to live information streams, we need agents to proactively filter data and inform users of relevant events as they occur. The generic cell could provide this capability.

For example, suppose a cell is linked to on-line stock prices that update a client's portfolio of stocks. The client might wish to know when one of the issues drops by 10 percent. By defining a formula in the cell that monitors an issue, a broker can cause price changes in that stock to trigger numeric analyses, audio, graphics, or even video. Because we can think of cells as independent agents, we can expect that soon a computer application will call clients on the telephone to inform them of changes in their portfolios—before a broker has the chance.

Applications for hypermedia spreadsheets

Spreadsheets already have the capability to access real-time stock quotes and news services. Time-shared systems and CD-ROM from several data vendors provide a reservoir of historical financial information. With the imminent realization that digital television will soon be available, we can imagine the "broadband telecommunications society of the future," in which a spreadsheet has access to network broadcast stories. The closed caption description goes directly into a cell. Other cells are defined with rules that look for specified topics in the closed caption. You might be working on an analysis of IBM when a window appears with a live broadcast on a merger announcement, as it happens.

If cells were generic objects that we could incorporate into documents, any computer environment could provide access to dynamic sets of data that are continually changed, analyzed, and updated. The integration of cells and multiple media on the desktop as hypermedia spreadsheets would serve several business applications well:

- Multimedia quarterly and annual reports
- Sales videos with pricing information
- Advertisements with circulation and demographic data

- CEOs' messages
- Executive information systems
- Training on a library of financial analyses
- Analyses of manufacturing facilities with efficiency calculations

Hypercalc

At the MIT Media Lab and the MIT Sloan School of Management, we set out to explore the implications of driving multimedia from a spreadsheet. We constructed Hypercalc, a prototype hypermedia spreadsheet, in the spring of 1990 by combining an authoring interface with a series of video device drivers written for Informix's Wingz, a graphical spreadsheet with a built-in scripting language. From within Hypercalc, an author can combine numbers, video, text, graphics, sound, and animation. A message in any of these media can be evoked by a numeric change in a cell. Hypercalc operates as a template for building multimedia applications.

To explore how business education might benefit from multimedia, we applied Hypercalc to business case study.

If cells were generic objects that we could incorporate into documents, any computer environment could provide access to dynamic sets of data that are continually changed, analyzed, and updated.

Business case study

Traditionally, business school faculty author the business case as a way of communicating management-related issues. The professor presents a particular corporation or industry from a viewpoint such as finance, marketing, or strategy. These studies typically have been distributed as textual documents with graphs and sometimes a spreadsheet.

Hypercase

The Hypercase model proposes a new framework for business case study.² Using our hypermedia spreadsheet Hypercalc, Hypercase structures an industry so that readers can gain a gestalt of it by interacting with the system. In defining how to analyze an industry and its competitors, Michael Porter identified two types of data: "published data and those gathered from interviews with industry participants and observers."³ The first three sections of Hypercase—Competitor Analysis, Industry Applications, and Industry Experts—contain such data. The fourth section contains a set of analytic tools that participants



Figure 2. The analytic toolbox.

can employ as they navigate through the material in the first three sections. Figure 2 shows the graphical interface for this analytic toolbox. By evaluating the case data, students learn the methods for industry research and analysis.

Within the Hypercase environment, the line between author and viewer often blurs. While the developer of the case is responsible for collecting source material, a simple set of tools allows another professor using the published materials to detail a particular issue or problem for students to track. Likewise, a student, while browsing information along a path, can interact with the material using a variety of analytic tools. These tools allow students to alter or add to the information and use the same set of tools used by the author and professor.

The first Hypercase prototype

In 1990 we developed "The Emerging Multimedia PC Industry," our first Hypercase prototype. We videotaped interviews, product demonstrations, and market applications and collected financial information on a range of multimedia PC hardware and software companies. We then edited the interviews and demonstrations and assembled them onto a videodisc that we integrated into a Hypercase. The project was developed on an Apple Macintosh II using a Mass Micro videographic card set and a two-screen display. The screenshots and examples in the following sections come from "The Emerging Multimedia PC Industry."

Analytic tools

The set of strategic tools in Hypercase draws directly from established management models such as the Five Forces Model,³ Growth Share Matrix (from the Boston Consulting Group), Life Cycle Diagram, and Value-Added Chain.⁴ Industry analysis, according to Porter, aims to unveil "the structural features of industries, the important forces causing them to change, and the strategic information . . . about competitors."³

As students browse through a Hypercase, they can use the strategic tools in the analytic toolbox to help crystalize their understanding of the industry's structure and of the relative strengths and weaknesses of its participants.

Hypercase tools

We can liken the information body in a Hypercase to a snapshot of an industry at a given moment. Though this information is arranged in a structured framework, we can imagine an infinite number of possible cross sections of the data world, each with relevance and each highlighting different concepts. This is the essence of traditional case study design. To teach a concept, a case writer will look to the real world to provide situational examples. The writer will "slice" through the set of data available to script a cohesive, bounded view of the real environment.

Using the tools in Hypercase, case writers, professors, stu-



Figure 3. Apple case using the Hypercase template.

dents, or other users can make their own slices of the Hypercase to illustrate concepts and pose new questions. A template helps guide case writers in producing an easily used case study with a standard Hypercase interface. This template appears in the screen shot of our sample case about Apple Computer (see Figure 3). Using the Hypercase tools, which include Segment-Builder, Sheet-Builder, Cell-Link, and Link-Tool, authors can enter data into a case template and link it to numbers, video segments, sound bytes, or text. Many of these multimedia programming constructs build upon the Elastic Tools, a hypercard toolset built for the Elastic Charles Project.⁵ Hypercase tools, whether graphical or menu driven, allow authors to concentrate on content rather than on the programming environment.

Segment Builder

Hypercase allows readers to browse video information by referencing segments. A video segment is defined as a sequence of linear video with a begin point, an end point, and a context that justifies its grouping. A video frame representing the segment—selected by the author—can be associated with a segment when the segment is defined; this pictorial icon later serves as a button to play the video segment.

Hypervideo

Hypertext, as coined by Ted Nelson,⁶ provides a metaphor for presenting textual information in a manner corresponding to the human process of association. When we read a body of text, certain words might cause us to reflect on related thoughts. Hypertext provides the ability to covertly include such links in a document by defining "hot words," which are indicated in bold face or by a unique color. Ideally, we would like the same link capability with any media format.

Video is a temporal medium. Therefore, we cannot define links solely between static segments. Over a time period within a segment, we must be able to define a link to another associated segment. In explaining their Elastic Tools, Brondmo and Davenport refer to this as "dynamic links in temporal media":

Links are used to associate content-related chunks with each other and ultimately to navigate between them. Some indication that a link has been associated with a chunk of information must present itself to the user. The user has the option of following the link or ignoring it.⁵

To provide this possibility within Hypercase, we created a linking tool that allows linking one segment to another segment, to multiple segments, or even to a spreadsheet. A link is defined within a segment by a start point, an end point, and an associated segment. During playback of a segment, an icon superimposed on the video indicates the existence of a link.

Cell-Link tools

Although many early multimedia applications allow users to interact with "sight, sound, and motion" (as defined by Phillip Schiller of Apple Computer in "The Emerging Multimedia PC Industry"), Hypercase puts a heavy emphasis on numbers. Not only can a user interact with numbers, but the numbers themselves can interact with other media. Using the Cell-Link tools, you can establish rules that link conditions to actions. For example, you can place threshold levels on a cell value that, when met, drive a relevant video segment to play on the monitor. The action does not impede the user from leaving the cell above the threshold level, but rather provides some "expert" advice or reminder regarding the decision.

In the case of "The Emerging Multimedia PC Industry," we implemented a brief exercise in which the student is to allocate staff between hardware and software development groups for a digital video R&D department. We defined a rule, or cell link, to activate when the ratio of software staff to hardware staff falls below 50 percent:

=IF(SoftwareRatio<50%, Play:Segment1,0)

If this rule becomes active, it triggers a video segment in which an industry representative (Kenneth Wirt) declares "... Multimedia without software is like a boat anchor. The days of developing hardware without software are over! Gone! Done! ..." In this instance, the educational impact of using a cell link is twofold. First, it offers students immediate feedback on their analy-

ses and assumptions. Second, using video as the feedback medium provides a dramatic and entertaining message.

Why multimedia for case study?

Why is multimedia an improvement over text? By looking specifically at case study, we were able to isolate several generic limitations of monomedia that a multimedia environment can improve upon. The following eight points outline some of the shortcomings of the traditional case study model.

1. Interpretation. Many managerial decisions must be based upon nonverbal, nonquantifiable cues. Human characteristics often make the difference. The decision maker wants to ask "Is this individual trustworthy?" "Is he articulate?" "Can he command the respect of others?" In current case study, direct state-

Hypercase puts a heavy emphasis on numbers. Not only can a user interact with numbers, but the numbers themselves can interact with other media.

ments of fact convey such information: The case writer, not the reader, performs the task of interpreting an individual's character traits.

2. Objectivity. Any publication or representation of reality must by definition carry some subjectivity from the author or artist. Even if the author provides unaltered data as collected, he or she must choose a search path in collecting that data and deciding which information to include. In a case study, we would like to get as close to an objective representation of the real world as possible to allow students to form their own views, untainted by the author's biases. Multimedia provides a better ability to publish objective data.

3. Linearity. Case study conveniently packages information for the reader. We begin on page one and read through to the end. Not only does this force the author's cognitive mapping on the reader, but the reading process has little likeness to the way we search for and collect information in the real world. When researching a topic, bits of information provide us with clues that can lead down multiple search paths. In the real world, the researcher must decide which path to follow and when.

4. Interactivity. The traditional case model provides information in a static medium. Case study need not be passive, unilateral communication. When preparing a case, you should be able to interact with it, testing assumptions and receiving comments.

5. Perspective and role play. No coin has just one side. Yet current case study often provides only a single perspective of a situation, thus leading readers to confuse perspective with fact. In addition, dialogue based on a single perspective usually boils

down to black-and-white conclusions. Because no two people see the world the same, issues are confronted and resolved by parties with different perspectives and goals. Ideally, we would like case study to have sufficient amounts of data to provide multiple perspectives on the same environment. Through role play of multiple perspectives, a more true-to-life discussion of an issue can evolve.

6. Scope and comprehension. Most case studies today have a narrowly defined scope. One reason for limited scope is the desire to highlight a particular concept. Another reason, which we feel drives the constraint, is limited time and overloaded memory capacity. The richer our active data set, the more meaningful is our analysis and discussion. Therefore, we would like to have an environment, such as multimedia, in which comprehension comes easily and information can be reviewed quickly.

7. Engagement. Disregarding the more philosophical problems with current case study method, we still see an equally grave shortcoming: They are often tedious. Students frequently read more than 100 cases in their first year of a masters program. Case study that engages and entertains, bringing readers into the material, might provide a compelling educational alternative.

8. From academics to applications. As you would expect from a methodology so far abstracted from reality, current case study as a framework does not present a smooth transition to real-world problems. If case study truly prepared us for the business world, we should be able to replace the limited data set with live feeds and use the same framework to analyze and understand actual environments.

Although we extracted these eight points from our considerations of case study, they might prove useful in determining the appropriateness of multimedia in other publications. If these categories are an important part of the message an author tries to convey, most probably a singular medium will not suffice. Multimedia, defined as a computational mixture of numbers, video, sound, text, and graphics, can be a viable solution.

Conclusions

Numbers are a medium that counts. By replicating the properties of a spreadsheet cell, we can ensure that the number returns to its dynamic context when mixed with other media. Providing computation within a range of document types is integral to the acceptance of hypermedia, because the business world works in the medium of numbers and spreadsheets. Incorporating the cell as the programmable unit for accessing multiple media types will ensure that multimedia becomes a form of communication for the mass market of computer users, not just a technology for a handful of developers.

The multimedia industry has referred to intelligent television as a receiver that will filter many real-world sources looking for data interesting to viewers. In the financial world, spreadsheet templates often define a layout and series of analyses for incoming numbers. If we stretch the analogy for a hypermedia spreadsheet, a template becomes an effective way to define a

personalized interface to the outside world. We could choose to have television stories on current events in one cell, live stock quotes in another, and a list of client activity in a third. Beneath the interface would be cell links defined to determine the content displayed.

Including numbers in an application as a dynamic medium enables the creation of an engaging, powerful, interactive document. Hypercalc and Hypercase provide a valuable platform on which to build business-related multimedia applications. The framework defined in the Hypercase environment can help managers analyze an unrelated industry, then extract the tools and methodologies needed to perform strategic analysis on their own company's behalf. Therefore, we think Hypermedia, as implemented in Hypercalc, can prove its effectiveness both as a learning environment and as a real-world management tool. □

Acknowledgments

Carlos Reategui was instrumental in the implementation of the Hypercase software. Thanks also to Dan Mayo for ideas brought up in conversation.

References

1. V. Bush, "As We May Think," *The Atlantic Monthly*, Vol. 176, July 1945, pp. 101-108.
2. J.D. Harber, "Hypercase: A Framework for Interactive Multimedia Case Study," Massachusetts Institute of Technology, Cambridge, Mass., 1990.
3. M.E. Porter, *Competitive Strategy Techniques for Analyzing Industries and Competitors*, The Free Press—A Division of Macmillan Publishing, New York, 1980, p. 368.
4. A.C. Hax and N.S. Majluf, *Strategic Management—An Integrative Perspective*, Prentice-Hall, Englewood Cliffs, N.J., 1984.
5. H.P. Brondmo and G. Davenport, "Creating and Viewing the Elastic Charles—a Hypermedia Journal," in *Hypertext: State of the Art*, R. McAlesse and C. Greene, eds., Intellect Ltd., Oxford, England, 1990. Also available from MIT Media Lab.
6. T. Nelson, *Computer Lib/Dream Machines*, published by the author, available from The Distributors, 7025 Michigan, South Bend, IN 46618.

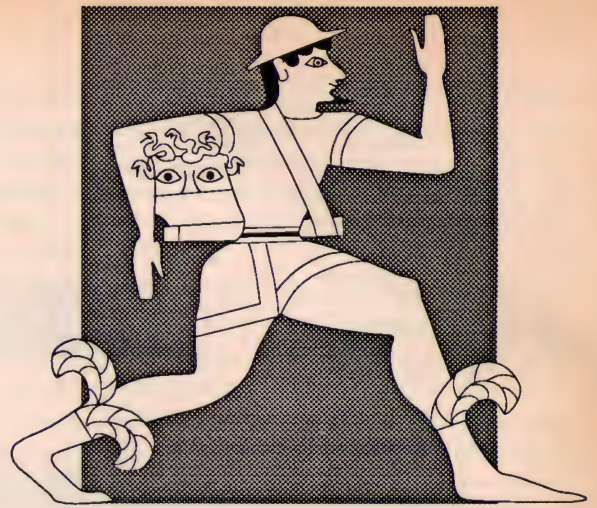


Glorianna Davenport is an assistant professor of media technology at Massachusetts Institute of Technology. She directs the Interactive Cinema Group at the Media Laboratory and has produced a range of interactive multimedia projects. Davenport received her master's degree in art history from Hunter College and her bachelor's degree in English from Mt. Holyoke College.



Jonathan D. Harber is currently president of Di Va Corporation, a Cambridge-based software firm that develops digital video applications and tools for the Macintosh computer. Harber received his MS degree at the MIT Sloan School of Management. He received his bachelor's degree with honors in cognitive science from Wesleyan University.

Readers may contact Davenport at MIT Media Lab, 20 Ames St., MIT E15-433, Cambridge, MA 02139 and Harber at Di Va Corp., PO Box 595, Kendall Square, Cambridge, MA 02142.



©1991 Perseus Project

Hypermedia and the Study of Ancient Culture

Gregory Crane
Harvard University

Allowing fluid movement among disparate data, hypermedia databases ideally suit the study of ancient cultures, where evidence ranges from literary texts to archaeological sites to individual objects.

Students of ancient Greek culture share with others an intense interest in tools allowing them to make new connections between the scattered aspects of their subject. Hypermedia databases allow them to cover a wider variety of evidence more thoroughly and make it less likely they will pursue the same text-limited questions as before. The long-term consequence will not simply be new answers, but new questions and research programs.

Students of ancient Greece have, as a body, integrated electronic databases as deeply into their basic research as any comparable group in the humanities. Classics have come to occupy an unusually prominent position in humanities computing generally, thanks to two complementary and far-sighted efforts. In 1972 planning began on the Thesaurus Linguae Graecae (TLG), a database project based at the University of California at Irvine and designed to place all ancient Greek literary texts on line. Then, during the mid-1970s, David Packard began developing computer systems to support this database and multilingual research (with an emphasis on Greek).

Classics departments began using minicomputer-based systems more than a decade ago and, since the mid-1980s, have had access to a personal computer specifically designed for their needs. Additional software made the massive TLG database accessible on general platforms such as the Macintosh and the PC. As of this writing, more than 500 departments and individuals use the TLG CD-ROM with its 300 Mbytes of Greek literature. Perhaps half of those actively engaged in research on ancient Greece have access to this resource, and that percentage is rising steadily.

Yet the very success of the TLG database poses problems for those inter-

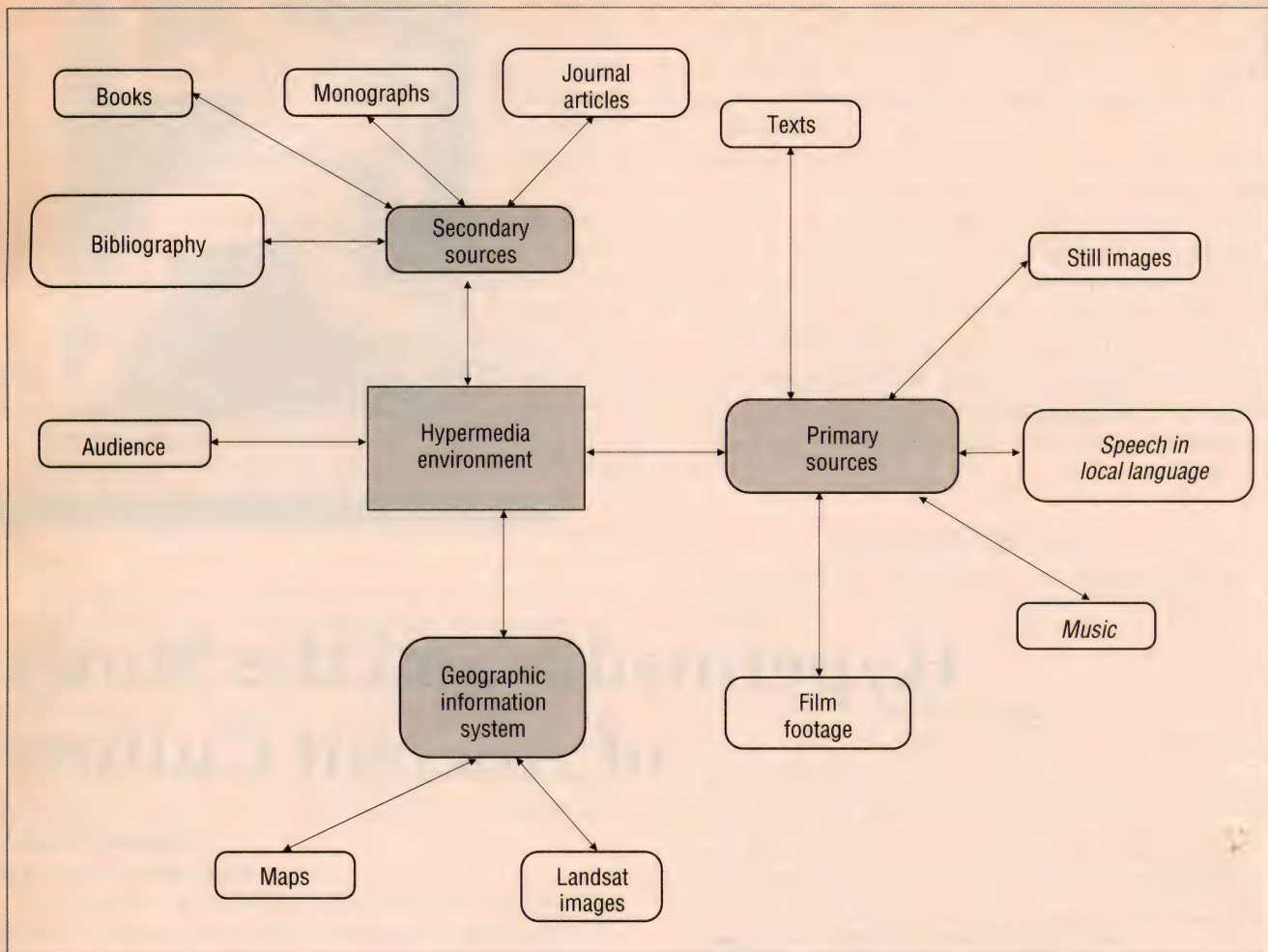


Figure 1. Simplified diagram of a hypermedia environment. Data types in italics are not included in the Perseus database. A second-order hierarchy such as that pictured is one way to organize our information.

ested in developing broadly associative hypermedia databases. (I use “hypermedia” rather than “multimedia” because it emphasizes associative browsing. Putting texts, images, sound, and motion video on line forces us to explore ways in which to link these different media. See Figure 1.) Scholars have enthusiastically embraced text retrieval, but the textual database mirrors the “monomedia” focus that has developed in many disciplines. Students of classical Greek literature, in particular, tend to work intensively with texts and pay comparatively little attention to other archaeological evidence.

In the past few years, for example, classicists have engaged in a debate as to the manner in which one kind of poetry was performed. Hundreds of painted Greek vases illustrate the performance of poetry and of music, and they provide one of our best sources of information on this topic. However, in the published articles the authors based their analyses entirely on the poetic text, citing an average of 10 separate passages per page to build their arguments. Of five recent articles, only one even alludes to the visual evidence—and that in a single footnote buttressing a secondary argument. Conversely, literary critics

considering archaeological material frequently restate standard opinions from published sources rather than reevaluating the evidence and building a new synthesis (as they would if working on texts).

The excessive emphasis on textual materials highlights both the short-term challenges and the long-term promise of hypermedia databases. The challenge will be for those who study ancient Greece to make good use of such databases. Literary criticism has spent more than 2,000 years developing topics best studied with textual evidence. In addition, despite the development of photography, printed images remain expensive and often of poor quality. Even specialists in classical art and archaeology find it unnecessarily time consuming to collect images illustrating a theme or topic. Those who can will often avoid the visual record altogether, selecting topics for which the visual record is not especially relevant or (as in the case of the poetry articles cited above) defining the problems in such a way as to dismiss the artistic and archaeological materials. Many consequently do not yet have the training or the background to use these materials.

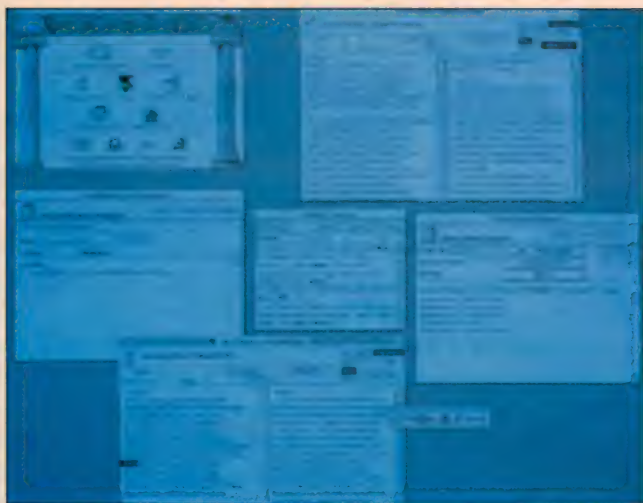


Figure 2. A screen from the Perseus hypermedia database showing a Greek and English text.

Nonetheless, few of the literary critics who focus so heavily on the text do so happily. Even the most quarrelsome factions in classics (and the humanities in general) would largely concur that literature must be studied in the context of the culture that produced it. The same forces that affected the form of a poem might have shaped the design of a temple or the scene painted on a Greek vase.

An experiment in hypermedia

The Perseus Project was conceived in the mid-1980s as both an experiment and a complement to the TLG. Where the TLG sought to collect all literary texts surviving from ancient Greece, the Perseus Project emphasized the earliest period of Greek civilization, roughly from the eighth century B.C. to the death of Alexander the Great in the late fourth century B.C.

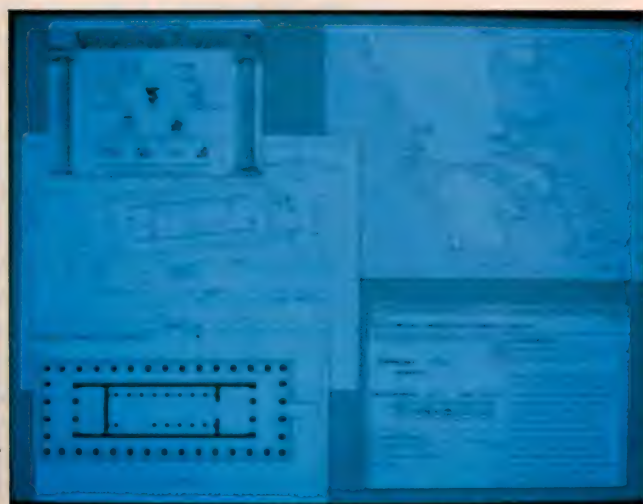


Figure 3. A screen from the Perseus hypermedia database showing an archaeological site.

However, the Perseus Project set out to collect as many kinds of data as possible from this 400 year period. We wanted to see what would happen to the study of antiquity if we could make not only more information, but more kinds of information, more readily accessible. We chose a hypermedia environment as the most flexible vehicle.

The Perseus hypermedia database includes source texts in Greek, English translations, essays, maps, drawings, and motion video and color still images representing Greek art and other archaeological materials¹⁻³ (see Figures 2-4). The database must support a wide variety of materials and, more importantly, serve a wide range of individuals, from the general reader to the advanced scholar. It must also stimulate inquiry into many different (and at times antagonistic) approaches.

A civilization—even one that, like classical Greece, has long vanished—poses questions for faculty in departments such as comparative literature, art history, archaeology, social anthropology, philosophy, political science, religion, and linguistics. Each of these disciplines requires a slightly different set of data.

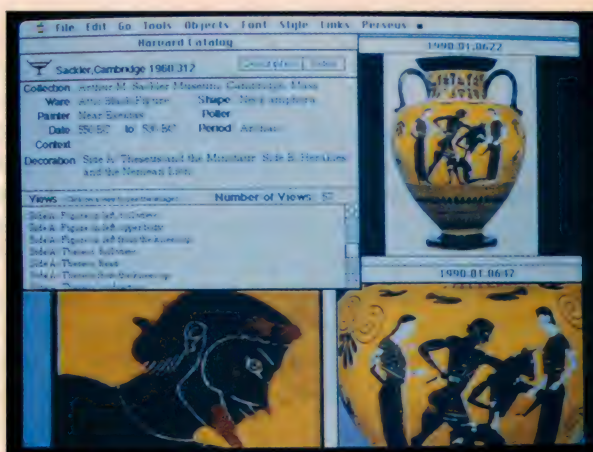


Figure 4. A screen from the Perseus hypermedia database showing a Greek vase.

For example, an art historian might study a painted Greek vase to determine which particular Athenian vase painter produced it and to further understand that painter's methods. In contrast, an archaeologist might have little or no interest in who painted the object, but rather in its discovery in a particular tomb in central Italy along with six other vases of the same shape.

The Perseus Project is equally a wholehearted exploration of one culture and a case study in how to represent culture in a hypermedia environment. It also probes the strengths and weaknesses of the new medium. Although the study of ancient Greece incorporates many different methodologies, our primary subject materials are easier to control than in many other fields. A textual database of 60 or 70 Mbytes allows us to include the majority of surviving literary texts, in both Greek original and English translation, as well as relevant later sources (such as sections of Plutarch).

Of course, enormous quantities of material remains exist. In just one category—painted Greek vases with scenes from mythology or daily life—we have roughly 80,000 whole or fragmentary vases. The *Princeton Encyclopedia of Classical Sites*, published in 1971, includes roughly 300 sites from the Greek mainland, many of which have produced vast amounts of information. Hundreds of thousands of coins and millions of artifacts have been recovered. Although we cannot expect to be comprehensive in this sphere, we have already collected roughly 18,000 images, including 1,000 drawings, plans, and maps, which will provide the most extensive collection of materials yet published.

We expect to distribute texts, drawings, database information, and other naturally digital forms of data on CD-ROM. We plan to make color images available on videodisc or on CD-ROM. In addition to providing broad coverage, we will select a few topics for intensive work. These will show the general audience what archaeological work entails and will illustrate to professional archaeologists what they can do in this new medium.

Representing Greek art

One particular experiment under way within the Perseus Project illustrates how a large hypermedia database can open up a subject to new audiences. Consider, for example, the way print publications represent and define our access to museum objects. Photography made possible the study of art as currently conceived, since few indeed can spend their lives in the agreeable practice of moving from one major museum to another. Nevertheless, even the modern catalog must compromise between logistical and economic factors. Books are bulky,

and reproducing pictures is expensive. The traditional publication must therefore minimize the number of illustrations, augmenting them with textual descriptions.

A full description of an object will refer to many details invisible in the photographs, details that might prove crucial to interpretation. For example, an overall view might obscure whether a scene's central figure wears a smile or a grimace. Moreover, black and white photographs simply cannot represent all the information communicated by color, but the expense of color illustrations means few books have many. Even painted Greek vases, which use color in a far more constrained manner than Western painting, exceed the capacities of halftones. Textual descriptions of black figure Attic vases regularly include a separate section such as "Added red: eyes, cloaks, shield rims of all Amazons." Because the illustration in the catalog will often wash out the differences in color, the author must describe them in text.

The possibilities and constraints of a large hypermedia database have led us to rethink the way we visually document objects. Whether we work with videodisc (and thus have access to 54,000 NTSC video images) or a CD-ROM filled with (for example) 10,000 compressed 640 × 480 24-bit color images, the new medium changes the quantity and quality of visual documentation that we can offer.

In photographing objects specifically for a hypermedia environment, we found that our standards for photographic coverage changed. The views we had previously selected were not the best possible views to illustrate the object in question, but represented the best we could do with four or five pictures. We now as a matter of course photograph each three-dimensional object from all sides, provide overviews of each major scene, and show details of all major figures. In gross terms, we regularly provide more images, such as shooting an average of 22 pictures for a Greek vase.

More subtly, we can better adapt our coverage to the object in question. For example, a printed catalog might provide between two and 10 views of the different objects that it contains. In the database, a small, lightly decorated piece might get only four or five overviews. On the other hand, more than 110 pictures illustrate a large, complex Greek vase in Harvard's Sackler Museum. The hypermedia environment allows us to tailor our photographic coverage far more closely to the needs of the object. Of course, our photographic coverage is itself a selection of views, and no careful viewer will find that we have covered everything of interest on every ob-

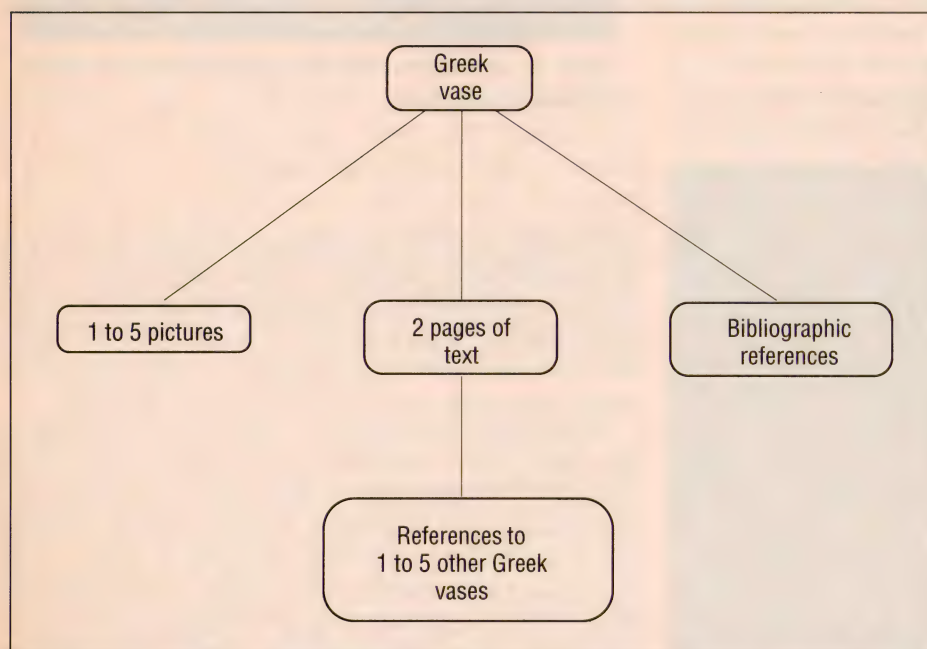


Figure 5. Diagram for a standard printed catalog entry for a Greek vase. Relatively few readers will have access to images of the other vases cited.

ject. Nevertheless, our audience will be able to see the objects in our database far more clearly than in print.

A hypermedia database does not simply affect the way we photograph an object. The problem for most viewers is context: Why do we think this vase was painted by a particular painter? What is the significance of the picture on the vase? Is it conventional or perhaps an odd variation on a particular theme? Such questions lend significance to an object, but are extremely hard to communicate in a printed publication. The only way to understand that a vase illustrates a conventional portrayal of Heracles fighting the Nemean lion is to see several other vases with the same scene, with each figure standing in the same relative position to the other. Conversely, stating that a vase illustrates a new approach to a particular scene will not suffice—readers need to see enough other conventional representations to properly appreciate the difference. Textual descriptions of visual information can lead the eye and point out significant features, but they cannot substitute for images of the objects to which they refer.

In the spring of 1990 Sarah Cormack and Mark Stansbury-O'Donnell, two graduate students in fine arts working for Perseus under the supervision of Susan Matheson (curator of the Yale Collection), created six catalog entries on Greek vases in the Yale Art Gallery. Composed from the start for distribution as part of the Perseus database, these entries came to assume a nontraditional form. Wherever possible, each statement about the object at Yale was documented by reference to an illustration of objects in other museums. Rather than presenting written judgments, we tried to create verbal links calling attention to the similarities between objects and directing readers to view the evidence for themselves. Each catalog entry included links to an average of 30 other objects.

Now if we contrast the electronic version with the printed catalog, we can begin to see the difference in form. A printed catalog entry for one art object might contain four or five pictures, a page or two of text, and a half dozen items of bibliography (see Figure 5). Its hypermedia counterpart would include a bit more text (perhaps three or four pages), because the 20 or so pictures of the main object would connect to an additional 30 objects, each (in theory) represented by its own 20 or so pictures. In the immediate future, anyone developing a hypermedia database will have to use black-and-white prints or color transparencies—the normal tools for print publication—to document most of the comparanda. A single CD-ROM will allow us to document roughly 90 objects, a figure comparable to

many printed catalogs (see Figure 6) if we make the following assumptions:

- We work exclusively with 640 × 480 24-bit color images compressed to roughly 50 Kbytes.
- We use a CD-ROM for storage.
- 22 images illustrate each primary object.
- Each object will have links to 30 additional objects for comparison.
- Each object chosen for comparison will be illustrated by three images (as opposed to the 20 views for primary objects).

The precise figures will, of course, vary. We might not choose to work exclusively with full-frame images, and we might not always use 30 comparanda for each object. Nevertheless, the general idea should be clear enough. If we conceive of a CD-ROM rather than a printed book as a means of publication, then we can treat the same problem (the documentation of interesting works of art) in a far more detailed manner. The hypermedia audience will see the individual objects far more clearly and will be much better prepared to understand the artistic context in which they were created.

The quantitative and qualitative change affects the range of questions hypermedia can stimulate. Readers cannot ask questions about evidence to which they do not have access, and access to primary material directly constrains the questions we choose to pursue. A richly documented publication such as the hypermedia museum catalog described above will, superficially at any rate, undermine the author's authority. Readers can see at least some of the evidence underlying the author's conclu-

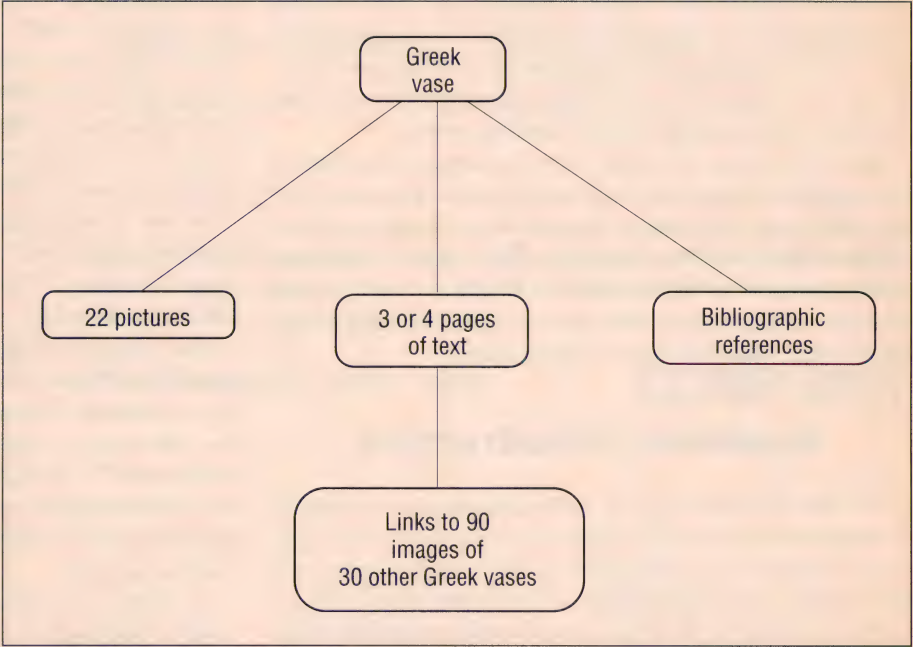


Figure 6. Diagram for a catalog entry for a Greek vase as created for the Perseus Project's hypermedia database. Note that the links in the hypermedia database are to images also accessible.

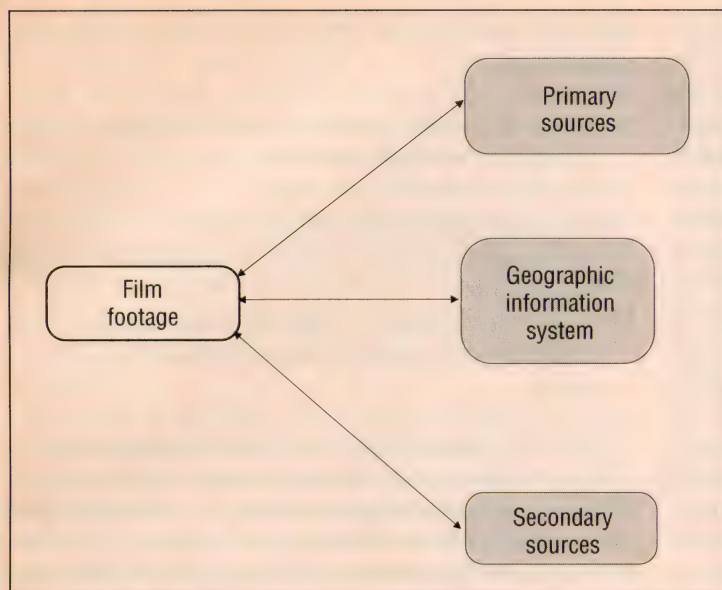


Figure 7. Video as part of a hypermedia environment. An electronic video stored as part of a hypermedia database can contain links to other kinds of materials. Moreover, other documents, such as more traditional textual discussions, can contain links to it.

sions, and a greater number of readers will be able to challenge those opinions.⁴ Nevertheless, art historians will be able to engage a far wider audience, and a greater number of researchers will be able to integrate artistic evidence into their work.

This hypermedia representation of art provides a single and, in some respects, fairly simple example. The main advance largely depends on a quantitative change in the amount of visual information we can publish. However, the simple ability to add more pictures allows us to represent objects in such a way that users can ask qualitatively different questions.

Our work so far on the Perseus Project suggests that over the coming years almost every tool used to study culture and society will be taken apart and redesigned to exploit the capabilities of hypermedia. Not only can hypermedia tools serve specialists more effectively, they can become accessible for the first time to many nonspecialists. Thus formerly esoteric subjects can play a larger role in the popular imagination.

Academic infrastructure

The narrow scope of printed books has a number of negative consequences for the way in which scholars disseminate their ideas. A single, linear book generally mixes basic evidence and the conclusions drawn from that evidence. The author arranges primary evidence in the sequence best suited to make particular points, choosing what to leave in and what to leave out.

The hierarchical, narrative form of a well-organized book leads the reader to scan a text in a particular order, giving the

author's conclusions more weight than is proper and making it more difficult for subsequent researchers to extend what has been done. Now, researchers returning from extensive fieldwork or archival study can, in a hypermedia database, publish thousands of pages of primary source material while more successfully keeping this material separate from their own conclusions. In addition, the author can include many images and even some limited motion video and sound to illustrate key points (see Figure 7). The reader in this environment is in a far better position to critique or to extend the author's conclusions.

Once authors can publish their ideas as a hypermedia database, they must alter their entire research process to take advantage of the new medium. They must more systematically collect audiovisual materials such as slides, recordings, and video. If the final publication is a CD-ROM-based database capable of holding 5,000 images, then taking 500 slides will not exploit the medium fully. Unfortunately, we do not yet train graduate students in the skills needed to make reasoned and intelligent decisions on collecting video footage that conveys a reasonable amount of information.

Benefits to come

The real benefits of hypermedia databases will come when they begin to proliferate and interact with one another. Authors cannot create every tool needed to describe a culture. Experts who concentrate on particular areas of New Guinea or Zaire, for example, might create detailed topographic maps of the villages in which they worked, but they will surely want access to a larger geographical database illustrating the climate and terrain of the general region. If more existing background material already exists, then it will be easier for authors to integrate their new material into a wider context and for readers to evaluate this new contribution.

Before different hypermedia documents created in different systems can interact, we need to establish basic standards of such interaction. The hypertext community has developed basic guidelines that might lead to a standard for the links within a hypertext, the so-called "Dexter" model.⁵ The Text Encoding Initiative, an international effort funded by the European Economic Community and the National Endowment for the Humanities, is developing standards for encoding textual documents.⁶ At the same time, we have begun to see standards for individual media such as still image and motion video compression. The bits and pieces of a complex standard are gradually taking form.

The problem for hypermedia databases extends beyond formalized languages. A hypermedia database that runs on many different systems must allow its audience to perform some useful task. If we simply replicate the tasks already performed using printed documents, we might well produce crippled and

poorly designed data that five or 10 years from now will not serve our needs or meet our intellectual demands.

The challenge is twofold. First, we must imagine how the technology can allow us to meet our existing goals at least as effectively as conventional printed tools. Second, we must determine what questions or tasks we have never before considered asking or performing that now become possible. While the first task provides a necessary point of departure, the second task creates an environment in which hypermedia will achieve its fullest impact. □

References

1. G. Crane, "Redefining the Book: Some Preliminary Problems," *Academic Computing*, Vol. 2, No. 5, Feb. 1988, pp. 6-11, 36-41.
2. G. Crane, "Formats of Print and Questions of Culture: Hypermedia and Scholarly Publishing," *Scholarly Publishing*, Vol. 21, No. 3, April 1990, pp. 131-155.
3. E. Mylonas et al., "The Perseus Project: Data in the Electronic Age," *Computing and the Classics*, University of Arizona Press, Tucson, 1991.
4. G. Crane "Composing Culture: the Authority of an Electronic Text," *Current Anthropology*, Vol. 32, June 1991.
5. F. Halasz and M. Schwartz, "The Dexter Hypertext Reference Model," *Proc. Hypertext Standardization Workshop*, 1990, Nat'l Institute of Standards, Gaithersberg, Md., pp. 95-133.
6. C.M. Sperberg-McQueen and L. Burnard, "Guidelines for the Encoding and Interchange of Machine-Readable Texts," 1990. Research sponsored by the Association for Computers and the Humanities; the Association for Computational Linguistics; the Association for Literacy and Linguistic Computing.



Gregory Crane is an associate professor of classics at Harvard University and the editor in chief of the Perseus Project, which is developing a large hypermedia database on ancient Greek civilization. He has published a book on

Homer and a number of articles on Greek literature, as well as numerous papers on the impact of hypermedia on the humanities.

Readers may contact the author at Dept. of the Classics, 319 Boylston Hall, Harvard University, Cambridge, MA 02138.

July 1991

Read with your earphones on!



Cover image: Wayne Lytle

The July issue on Computer-Generated Music lets you...

Sample the standards!

- Mozart, Bach, and Brahms
- Jazz
- Folk music, and
- More

Encounter new approaches!

- Systems for machine composition
- Composition languages
- Pentatonic music, and
- Music of the planets

With your own 60-minute CD or cassette!

Digitally recorded music on compact disk or Dolby B audiocassette presents project results as you read the articles. Order today so you can read *COMPUTER* with your earphones on.

Music/magazine order form

Please print

Name _____
Address _____
City _____
State/ZIP _____
Country _____
Member No. _____
(required for magazine member price)

COMPUTER

	Qty.	Total
CD: \$16*		
Tape: \$12*		
Single issue price (member): \$10		
Nonmember: \$20		
CA residents add 6.5% sales tax; Canada residents add 7% GST		
TOTAL (enclose full payment):		

Mail to
IEEE Computer Society Order Dept.
PO Box 3014
Los Alamitos, CA 90720-1264

* Includes \$4.00 postage and handling.



Developing Multimedia Applications

Sorel Reisman
California State University, Fullerton

Experience with videodisc-based multimedia projects can provide development skill sets that you can use to produce your own multimedia applications. A multiphased approach shows you how.

Multimedia computing is not new, having existed as a discipline for almost 30 years. This article reviews the history of that discipline and provides a structured set of recommendations for those wishing to develop their own multimedia applications. These pragmatic, operational recommendations are based on three multimedia projects successfully completed inexpensively at California State University, Fullerton. Although the projects all used interactive laser disc systems and products, the experience gained transfers directly to newer, emerging multimedia technologies.

History lesson

End-user computing is more than 15 years old, but recent improvements in hardware technology have impelled dramatic changes to the interface between the computer and end user. The most notable example is the growing popularity of graphical user interfaces (GUIs) such as Apple's Macintosh desktop, Microsoft's Windows, and IBM's Presentation Manager. An extension of the interest in GUI's appears in the increased attention to interactive multimedia.¹ In the simplest sense, multimedia computing enhances a GUI when a GUI-based application responds to user input by

generating audio and full motion video output. With today's technologies the potential to deliver this kind of artificial reality depends on a powerful computer's ability to quickly access and manipulate data stored on a very large capacity secondary storage device.²

Neither the term multimedia nor the concepts that it embraces are new. Educators and trainers have used multimedia equipment for decades, including projectors and recorders for 16-mm film, 8-mm film loops, 35-mm slides, and audio tape. In a computer sense, instructors used this equipment in an off-line mode to enhance and supplement instruction. IBM was the first company to put multimedia equipment on line with the IBM 1500 System. Promoted in the mid 1960s, the IBM 1500 consisted of special-purpose workstations controlled by a conventional, time-shared IBM 1800 minicomputer. Analog audio and slide projection devices locally attached to each workstation provided users with individualized, multimedia, computer-based learning.

Although only a small number of 1500s shipped, the product served as a model for two other prominent, competing, National Science Foundation-funded systems: Plato, a mainframe-based system developed at the University of Illinois, and Ticcit, developed by Mitre Corporation.³ A limitation of these systems was that they used unreliable and slow mechanical devices for their audio and video storage sources. Despite this, and despite their lack of market success, through the 1970s these systems served as models for the ongoing development of other multimedia computing systems.^{4,5}

By the early 1980s, when more reliable and cost-effective optical-based products became feasible, application developers had learned a great deal about the do's and don'ts of developing multimedia applications. During the 1980s, as developers standardized on the laser disc as a primary audiovisual storage medium, the experience base for multimedia application development surged. Concurrently, developments in programming tools (specifically, multimedia authoring systems), more widespread availability of audiovisual production tools (such as home video products), and improvements in disc manufacturing processes facilitated the creation of new multimedia applications.

Today we see a great deal of speculation and confusion regarding the likely (and inevitable) successor to laser discs. The alphabet-soup list consists of contending products and technologies such as DVI (digital video interactive), CD-I (compact disc interactive), CD-ROM-XA (compact disc, read only memory, extended audio), and CDTV (compact disc television). Others will likely join this list. Examination reveals consistent features that application developers can count on in the 1990s.⁶ All use optically based, randomly accessible, digital storage media and access devices. Moreover, they all capitalize on the trend toward decreased physical size with concomitant increases in storage densities at substantially reduced per-megabyte costs. These features are essential because of multimedia applications' dependency on fast access to high storage capacity media and devices.

Because of the similarities among these devices and media, application developers should build upon their earlier experiences to avoid costly application development errors made in the past. More importantly, they should learn how to expeditiously and cost-effectively develop the kinds of multimedia applications that will prove useful to the broadest range of end users.

Application developers should learn how to expeditiously and cost-effectively develop the kinds of multimedia applications that will prove useful to the broadest range of end users.

CSUF's approach

Since 1986 California State University, Fullerton (CSUF) has been involved in the development of videodisc-based multimedia applications.⁷ The results have been extremely satisfying, allowing the university to make great progress using multimedia in a wide range of subject areas. Most of the projects concerned the production of new audio and video materials; all were completed with minimal resources in less than six months. This section describes three of those projects and provides a set of recommendations for new and inexperienced multimedia application developers. Even though the CSUF applications were laser disc-based, the lessons learned are relevant to the next generation of multimedia applications.

Videotape transfer to disc

The purpose of this project was to explore the mechanics of creating a videodisc and to use it in a computer-controlled environment. We obtained permission from a national textbook publisher to master to videodisc eight copyrighted and professionally produced 1/2-inch VHS videotapes. To explore the random accessibility of the discs under computer control, we developed the computer program "The Authoring and Presentation System" (TAPS) for use on the IBM Infowindow System. That platform had emerged in the 1980s as the de facto laser disc system standard.

TAPS allows a non-computer-literate presenter to easily specify videodisc locations (frames) that the computer causes to be sought and played. TAPS was used with one of the discs described above to present subject matter discussed in management information systems lectures. The lecturer could select menu items from the Infowindow System screen and, via an attached large-screen video projector, display selected material to the class.



Figure 1. The opening screen of The Authoring and Presentation System's (TAPS) authoring module.

Figure 1 shows the opening screen from TAPS' authoring module. Figure 2 displays video on global telecommunications from TAPS' presentation module. Here, an instructor has interrupted the video during a lecture.

Assembling existing film footage

The second project, the creation of a campus directory videodisc, was more ambitious. The videodisc allowed users to use an IBM Infowindow System to interactively enquire about and explore campus activities and facilities. Figure 3 shows the main menu options, which a user would select by touch.

This project had a very low budget of \$3,000. Consequently, we had to use existing film/video footage shot for campus outreach (promotional) purposes. We selected, edited, and assembled on videotape those vignettes appropriate to a campus directory videodisc.

Shooting new footage

The third project, even more complex, concerned the development of a disc that could be used for group decision processes.⁸ GDP is a videodisc-based group decision support system in which a prerecorded human facilitator guides project team members toward consensus on an assigned task. Figure 4 shows GDP main menu items, while Figure 5 shows the GDP human facilitator.

Although the subject matter of the campus directory disc and GDP differed substantially, the projects still had much in common. Both required the development of a script around which the computer and audiovisual functions could be planned, the development of computer-based logic, and the generation of audio and video materials that could be mastered onto disc.

Scripting of the applications forced the developers to conceptualize and precisely document the logic flow of each applica-



Figure 2. An example of a TAPS' presentation module displaying video (on global telecommunications) interrupted by an instructor during a lecture.

tion. This phase resembled the analysis and design phases typical of the development of complex information systems. The logic development (computer programming) was done using IBM's Infowindow Presentation System. Novice programmers programmed, debugged, and tested each application in less than five months.

The GDP project benefited from the success of the campus directory disc, consequently receiving funds to produce new audio and video footage. Using faculty and students from CSUF's drama department, the developers shot approximately 15 minutes of new footage and later edited it to remove material that would not be accessed by program logic.



Figure 3. Campus Directory Disc showing touch-selectable main menu options.

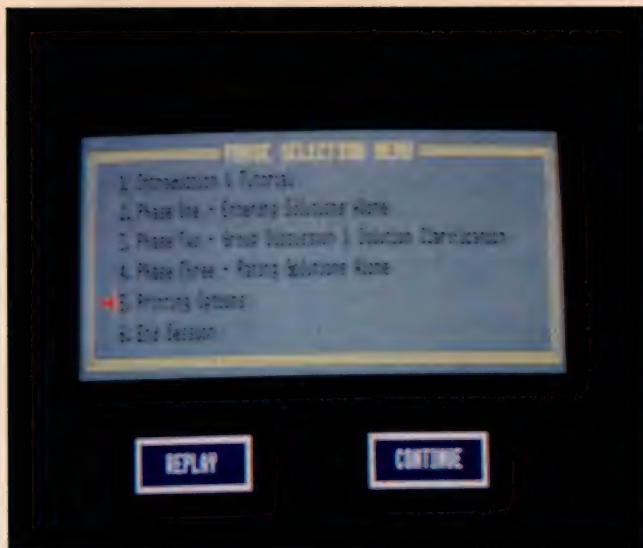


Figure 4. Group Decision Process (GDP) showing multiphase main menu items.



Figure 5. GDP screen of the human facilitator (at rest!).

In both projects the resultant video masters were recorded on conventional 1/2-inch VHS videotape. Although most professionals disdain the low quality of this format and urge disc producers to use more professional formats—preferably 1 inch—such facilities were simply not available for these projects. Additionally, the cost to use off-campus “professional” services exceeded the available budget.

Recommended projects

Drawing from our experience at CSUF, I can specify a number of principles and recommendations that those wishing to develop multimedia applications should consider. The guidelines appear as a series of projects in Table 1, with details in the next section. I recommend pursuing these activities in the order shown. Note that although the CSUF activities were based on three different projects, this table lists four. For those unfamiliar with script writing, that process is sufficiently complex that it warrants special attention in a separate project.

Remember that each of the descriptions is conceptual in nature, requiring significantly more work than described. However, as we learned at CSUF, these “go slow” projects can save time and energy, and require a minimal budget.

Project 1: Disc manufacturing

If your application will be based on either a laser disc or a CD-ROM, define a project that provides disc mastering experience. For a laser disc-based application, select for transfer to disc an existing 1/2-inch VHS or Beta videotape segment less than 30 minutes long. If you are building a CD-ROM-based application, select an existing digital database that contains the kinds of materials you will eventually want. For example, if your application is to be hypertext-based, you will have to build

the indexed text (and perhaps image) files that your search program will access.

In any case, select source materials with meaningful and relevant content useful for end users and testable in a real environment. Obtain copyright releases for materials you use.

A well-known shortcoming of laser discs and CD-ROMs is their read-only nature. Consequently, once discs have been pressed in quantity, you cannot change errors. Disc manufacturers provide “check” discs for developers to test their laser disc application before volume production. These discs are fully featured and fully compatible, but of lower quality than mass-

Table 1. Summary of recommended projects.

Project	Goal	Key Activity
1	Gain disc manufacturing experience	Transfer existing audiovisual and/or database to disc
2	Work with multimedia application development software	Repurpose an existing disc
3	Work with existing audiovisual materials to develop a new multimedia application	Create new script; edit existing footage for new disc; create new audio; program the application
4	Develop new audiovisual materials for a new multimedia application	Create new script; shoot and edit new audiovisual footage for new disc; program the application

produced laser discs. Most disc pressing companies manufacture a check disc for as little as \$300, substantially less than the \$2,000 start-up cost for volume disc pressing. The process of assembling and submitting material to manufacture a check disc is identical to that for volume production. I strongly recommend that every laser disc-based application be tested with a check disc.

Similar services are not available for CD-ROMs. The reason for this may be that the start-up cost to manufacture a single CD-ROM for a check disc is the same as the cost for volume manufacturing.

If you wish to avoid creating a CD-ROM with errors, you can emulate CD-ROM features using an approximately 700-Mbyte (magnetic) hard disk. This is functionally equivalent to using a CD-ROM, but access speeds will be considerably faster than for a CD-ROM drive. Note that text-based CD-ROM databases are usually generated by application programs (available from many vendors) that create an index file potentially as large as the data file itself. This implies that the maximum effective capacity of the CD-ROM containing the application might reduce to 300-350 Mbytes. Once the materials are acceptable, you can back them up onto a submission medium, usually a 9 track tape.

An alternative to testing with a hard disk is to use a WORM drive (write once read many). Although WORM drives cost approximately twice as much as 700-Mbyte hard drives, a WORM cartridge costs about \$100 and is removable from the drive. Some disc pressers accept WORM cartridges as input media for CD-ROM disc pressing. However, you cannot change data written to a WORM cartridge if you discover errors during testing.

Regardless of the target optical medium, it is essential to verify the media and formats that disc manufacturers accept as source materials for manufacturing. A project such as this is simple to undertake. It can provide experience in (1) dealing with disc manufacturers, (2) testing discs, (3) verifying the quality of check discs, and (4) understanding the realities of project scheduling.

Project 2: Repurposing

Define a project that provides experience with an authoring system or, for example, a hypermedia retrieval system, either of which utilizes the special features of multimedia applications. The essence of these applications is a disc's ability to be used as a randomly accessible multimedia database. Disc-based databases integrated into multimedia computing systems provide developers with an incredible palette of novel features.

For example, some of the more basic but still unique features of multimedia are forward and backward slow-motion video play at variable speeds, single frame motion, and multiple track audio. Even more unique features include the ability to display graphics images over video, to generate graphics animation over still video images, to experiment with special effects such as fading into or out of video scenes, or to graphically "window" portions of a video screen.

The trend toward digitization of conventional video provides a dimension of special effects that have only been available with expensive film or tape editing equipment. These effects vary from system to system.

While the implementation of such functions might seem complex, most commercially available laser disc authoring systems have been designed to facilitate their use.⁹ As newer CD-ROM-

In many ways repurposing presents a greater intellectual challenge than does the creation of original audio and/or video material.

based standards emerge, so have high-level authoring systems to support digital special effects built into the hardware. For example, the features of Intel's DVI hardware can be used by authoring systems from Ceit Systems (Authology: Multimedia) and Network Technology Corporation (Mediascript), and by paint programs from Time Arts (Lumena). Complexity lies not in the computer programming, but in using these functions creatively to develop unique multimedia applications. The ability to do this can only be acquired through hands-on experience with the hardware and software that provide these functions.

One way to develop this expertise is to formulate a project that requires the "repurposing" of an existing laser disc or CD-ROM. Repurposing refers to the development of computer programs that use an existing disc (as an audiovisual database) for an application different from the one for which the disc was originally intended. Repurposing allows the developer to focus on the nature of the computer-based authoring system rather than on the development of audiovisual materials. In many ways repurposing presents a greater intellectual challenge than does the creation of original audio and/or video material. The intricacies of adapting existing audio and video to repurposing objectives can teach computer programming techniques usable in developing original scripts in subsequent projects.

Project 3: Script writing

Define a project that provides experience in creating a multimedia application based on an original idea. This project should not entail the shooting of new footage. Instead, select existing film/video footage and (after the fact) write a plan (script) for its use in an interactive, multimedia environment. Obtain copyright releases for audio and video material you want to use. Plan for the script to include original narration to accompany the video. Because most new CD-ROM-based products provide multiple "tracks" of audio, consider developing a multipurposed application by scripting additional audio segments to accompany one or more of the conventional text, graphics, or video segments. You might use different audio segments for multilingual versions of the application. Or use two narratives

with a video segment, one perhaps as narrative for introductory users and the other for more sophisticated users.

From the disc manufacturer determine any special requirements for assembling the master tape. For example, tapes with more than 30 minutes of video cannot usually be used to manufacture an interactive laser disc. A standard, randomly accessible single-sided (CAV) disc contains a maximum of 30 minutes of video or up to 54,000 still images or a mixture of these. Disc manufacturers require adherence to certain guidelines for complex mixtures of audio and video materials. For CD-ROMs, depending on which of the many file formats you use, other limitations may exist.

This project will familiarize you with scripting, editing, and assembling the submission medium master planned for disc mastering. You will also become familiar with using special features of multimedia technology. If you decide to create original audio for the existing video, you will be introduced in a very gentle fashion to some of the issues involved in creating new material. It is much easier to create an audio narrative and edit it onto existing video than to shoot original video and audio at the same time. Choosing this simpler strategy provides production experience that will be valuable in subsequent projects.

Project 4: Integration of expertise

The activities of all the previous projects can produce unique results when applied to a single project. Because of the complexity of integrating newly acquired skills into a unique multimedia application, it is difficult to conceptualize the complete process the first time. Therefore, I strongly recommend that you attend one of the many commercially offered courses describing the totality of such a development project. The perspectives of these courses differ, depending on whether they are offered by a computer company or a video company (like Apple Computer or Sony Corp.), but most will include material concerning project planning methodologies that are essential for the successful completion of a complex multimedia application project.

The development of such an application is a team effort and requires the supervision of a project leader responsible for budgets and schedules. The team must consist of a subject-matter expert responsible for the content of the script; someone with video expertise responsible for the form of the script, as well as for all aspects of the video production; and someone with computer expertise responsible for computer programming, media integration, and application testing.

The first project should be modest and not scripted to require exotic filming conditions, nor should it depend on the most extravagant features of the computer-based multimedia system. All the team members should develop the script and iteratively revise it until finalized. Once the final script has been agreed upon, development can begin.

Various development activities can take place in parallel. For example, after the script has been finalized, computer programming can begin immediately and all the logic can be developed, tested, and debugged while the video production takes place.

The subject-matter expert should be available throughout all aspects of the video production to ensure error-free recording of the audio and video material. Unlike flawed computer programs, flawed multimedia programs are difficult to correct or to work around after editing is complete. When the audio and video have been edited, create a test audiovisual database (with a check disc, hard disk, or WORM disc), integrate it with the computer logic, and test the application. In all likelihood you will find many serious errors that will force you to revise some of the original script.

For example, if you find errors in the audio, you might have to present that content with displayed text. If you find errors in the video, you might have to play audio while displaying a newly created text, graphics, or still video image in place of the motion video. If your application is based on a CD-ROM with "conventionally formatted" text or graphics files, and if some of those read-only files are in error or damaged, you might have to patch your control software to access undamaged files resident on a supplemental hard or floppy disk.

Although your final product may fall short of the one you originally envisaged, you might decide that it is good enough. Alternatively, you might decide that reshooting or re-editing a small part of the audio or video is justified and would result in the "masterpiece" that you had been working towards in this and all your previous projects. □

References

1. S. Aubron and K. Hooper, *Interactive Multimedia*, Microsoft Press, Redmond, Wash., 1988.
2. J.D. Foley, "Interfaces for Advanced Computing," *Scientific American*, Vol. 257, No. 4, Oct. 1987, pp. 127-135.
3. E.J. Anastasio, "Evaluation of the Educational Effectiveness of PLATO and TICCIT," in *Frontiers in Education*, J.M. Biedenban and L.P. Graywon, eds., IEEE Press, N.Y., 1973, pp. 382-387.
4. S. Reisman and D.B. Martin, "Television Technology in Computer Based Training," *IEEE Conf. Digest*, Toronto, Sept. 1977, pp. 38-40.
5. S. Reisman, "Research in Multimedia Computer Assisted Instruction," paper presented at Annual Meeting of the American Educational Research Association, Toronto, March, 1978.
6. S. Oberlin and J. Cox, *Microsoft 1989-1990 CD-ROM Yearbook*, Microsoft Press, Redmond, Wash., 1989.
7. S. Reisman, "Interactive Videodisc at Cal State Fullerton," *Proc. 7th Conf. on Interactive Instruction Delivery*, Feb. 1989, Orlando, Florida, pp. 108-111.
8. S. Reisman et al., "Interactive Videodisc for Nominal Group Decision Making," paper presented at CORSE/DIMS/ROSA Joint National Meeting, Vancouver, BC, May, 1989.
9. "Authoring Systems, Programs, Languages," *Instruction Delivery Systems*, Vol. 3, No. 2, March/April 1989, pp. 10-15.



Sorel Reisman is an associate professor of management information systems in the business school at California State University, Fullerton, where he founded and now codirects the Interactive Videodisc Learning Center.

Reisman has been a member of the editorial board of *IEEE Software* and held the position of products review editor. He is currently editor of the new department, "End User," which he originated. He is also a contributing product reviewer for *Computer*.

Readers can contact Reisman at Dept. of Management Science, California State University, Fullerton, Fullerton, CA 92634.

Dolphins are currently threatened by abusive fishing techniques.



Learning through Multimedia

Roy D. Pea

Institute for Research on Learning

Teaching both new subject matter and new technology, Multimedia Works encourages students to create multimedia compositions that explain, tell a story, or persuade.

As multimedia technology proliferates, we must explore some central learning questions. Whereas movies separate directors and viewers, the technology of interactive multimedia can unite creators and viewers in much the same way as the technology of writing unites writers and readers. Society might come to regard multimedia literacy as essential as writing is today. Anticipating this future, the Multimedia Works Project focuses on enabling students to create and use multimedia documents. Our major challenge has been to design tools that young learners can use to create multimedia communications. These tools enable students to become multimedia composers, not just consumers. (I use the term “composer” rather than “author” because it correctly implies that creating multimedia involves designing an arrangement of parts.)

Learners need to control computer-controlled multimedia to communicate their own understanding of information. Multimedia objects (text, pictures, video clips, and so forth) are important building blocks for developing and conveying a student’s understanding for several reasons:

1. Multimedia communication is similar to face-to-face communication.
2. Multimedia is less restricted than written text. Many people come to understand text better with broader media support for its interpretation.
3. Multimedia can place abstract concepts in a specific context (for example, refraction in physics might be depicted in a film of lens and light behavior).
4. Multimedia allows for individual differences in preferred sensory channels for learning.
5. Multimedia lets you coordinate diverse external representations (with distinctive strengths) for different perspectives.

Creating multimedia composition tools for young learners is no small challenge. It means more than letting learners make choices among predefined options. Consider how students go about scholarly research. They develop a critical synthesis to summarize literature or film, craft a compelling argument, and carefully weigh pros and cons. In short, they construct meaning from cultural artifacts, a process similar to the composition of multimedia interpretations. Unless students start composing their own multimedia interpretations, multimedia “educational materials” will be “delivered” to students, just as books are now. We might find the 21st century recapitulating extant problems with texts in education, which promote an “authority-centered” epistemology.

Today’s book report or research paper is one place to start thinking about multimedia composing. With these assignments, students typically underuse research resources, are limited to written text as their final product, and lack opportunities for presenting and discussing that product with a critical community (that is, teachers and peers). When we extend these assignments to connect with large, user-extensible multimedia databases, collaborative activities, and interactive presentations to reactive audiences, then we come closer to a culturally-based learning activity centered on students’ constructing and communicating their understanding.

Most of multimedia computing in education or business now is what we call *chained multimedia* rather than *integrated multimedia*. In chained multimedia, the communication consists of the temporal sequencing of experientially distinct multimedia objects, like an interruptible slide show that accompanies text. A screen of text might contain a button to press that brings up a screen graphic or video sequence on another monitor, which might return to the text, where another button plays an animation, and so forth. Integrated multimedia makes spatial juxtaposition and temporal overlapping of diverse media available. Such integration is evident in the classic audiovisual fusion of sound and image in video and film.¹

It’s difficult to create even linear integrated multimedia, and interactive multimedia composition is even more complex. We decided to first resolve problems in linear multimedia composing, then build toward interactive, integrated multimedia. (Hypermedia networks of information, interactive books, and

“guided tours” with optional paths raise even more difficult questions.²)

Designing composing tools

Multimedia composing is a complex cognitive task. Multimedia Works’ goal was to support the user’s choreographic sequencing of media elements (video clips, text displays, photos, graphics, sound clips, animations) in space and time through a task-transparent interface. Until recently there were so many technical barriers to such choreography that we could not easily see the conceptual issues of integrated multimedia design.

Composing multimedia with existing tools is obviously more complex than writing. But various cognitive technologies for writing are available to support the writer’s work: planning and goal-setting, generating and organizing ideas and text, goal

Multimedia is less restricted than written text. Many people come to understand text better with broader media support for its interpretation.

progress monitoring, text evaluation, and revision.³ Researchers have given far less attention to the multidocument search and synthesis activities involved in writing. All these activities are key to multimedia composition.

Multimedia composing resembles film editing, animation, and/or choreography, except that you also want to have computational control over the objects and events you design. Such computational control includes where, when, and how long the objects should be displayed on the screen(s). These issues are largely treated as defaults in writing, but they are important variable properties in multimedia.

We knew that multimedia composing could be made much simpler by designing more congenial representations for creating, transforming, controlling, and viewing media elements. Multimedia composers need scores for coordinating displays of sound, image, and text over the space of display screens and the time of “playing” a composition.⁴ These scores should provide alternate views on the structure of multimedia compositions, that is, making patterns of media dominance or co-occurrence more visible.

In another article, Mills and I⁵ stated that multimedia composing tasks involve operations—cognitive and physical—on a number of hierarchically nested units. For example, a multimedia composition might constitute a story composed of episodes and events. Events consist of dynamic arrangements of scenes that express object relations. We found this scheme not wholly satisfying in practice. It’s far more important to consider the composing medium’s material substrate (properties of file



Figure 1. The Media Space interface for multimedia database browsing.

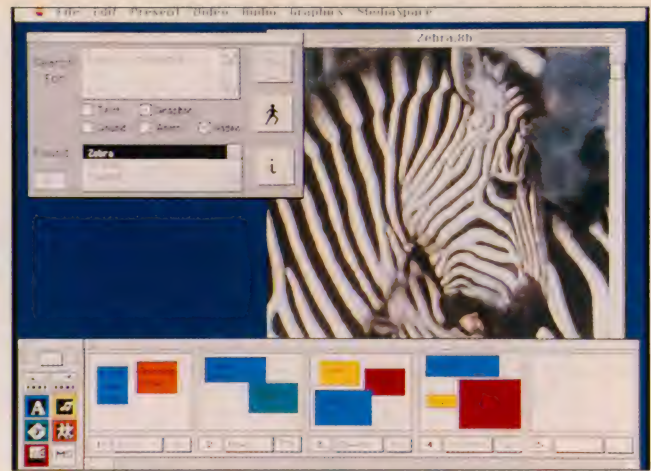


Figure 2. The Media Space interface for multimedia database search.

types, screen displays, and so forth) in defining the units of the hierarchy.

Multimedia Works software

The Multimedia Works software environment consists of three programs: Media Space, a multimedia database and research program; Multimedia Works Composer, a storyboarding and presentation program; and Video Light Table, a direct manipulation video-clip editor.

We initially designed Media Space and Multimedia Works Composer in Hypercard, then ported them to the Supercard environment to take advantage of the color and multiple windowing capabilities it offered. Supercard features, such as menu and window type controls and resource handling, greatly simplified the software development process and improved the user interface.

Media Space

Media Space currently consists of more than 500 text, graphic, sound, animation, and video elements that can be searched using key words or browsed using a visual index. This database occupies approximately 30 Mbytes on a local hard disk. Media Space is made up of a variety of search and editing tools coupled with a multimedia database of digitized elements. The media elements serve as a resource for students' explorations of environmental topics such as climatic change, pollution, and endangered species. Students can browse through a variety of color pictures and black-and-white graphics, listen to samples of digitized sounds, read a variety of topical news stories from diverse journalistic and reference sources, or view a more limited selection of digital animations and video clips (analog sources).

Students use Media Space in Multimedia Works both as a research tool to browse (as in Figure 1) or systematically to search (Figure 2) a wide array of photographic images, graph-

ics, selected readings, and digitized sounds. They also use it as a source of media in composing (with Multimedia Works Composer, represented in the storyboard at the bottom of Figures 1 and 2). Students access Media Space from the media palette (Figure 1).

Users can add media to Media Space with relative ease. Text can be typed in, downloaded electronically (using Compuserve or Dialog, for example), or taken from typewritten or typeset material using a scanner and text recognition software. Users can use a color scanner to digitize images, import them from other computer graphics applications, or digitize them from a video source. Using Farallon's Soundedit, hypermedia composers can also digitize sounds (which we have found easy enough for children to do).

Multimedia Works Composer

The heart of Multimedia Works, the Composer, is a powerful, easy-to-use multimedia composition and presentation tool. Multimedia Works Composer presents the user with a storyboard of 15 panels or media "scenes." A Multimedia Works composition consists of a number of such scenes to be played over time. Within the scenes, icons that represent media displays of different media elements are arranged spatially (Figure 3). A live storyboard of movable and editable media icons represents the document "filmstrip" made up of these scenes.

With tools to arrange media elements spatially and temporally, the Composer helps the user design the composition. Typically, students begin a new composition with a blank storyboard, then add different media elements to it by searching the Media Space database for appropriate elements. They also typically create their own elements with text and paint/drawing tools native to the Multimedia Works environment. They can create their own video clips (with the Video Light Table, described later), sound clips (with Soundedit), and animation (Macromind Director). Students can select these



Figure 3. The Multimedia Works Composer storyboard for design of multimedia documents.

tools from a menu within the Composer and then import the media files produced into their composition.

Students create media elements by clicking on one of the five media type icons on the media-creation palette (shown in Figure 3). For example, a user creating a text element is led through a dialog box process to (1) name the text card; (2) choose its scene; (3) create the text; and (4) change the text style or size.

Using the controls in the top left of the media palette, users can add media elements or delete them from a composition by placing them in the storyboard's numbered scenes. Once placed, the elements are represented by small icons called Media Boxes, which are color-coded to distinguish the media types. Each media element's name appears on its colored Media Box. A scene can be viewed by clicking its numbered button, which highlights the border of the selected scene. Each media element in that scene is then displayed behind the foreground of the live storyboard (see Figure 3). Users can reposition, crop, or enlarge the media elements in a scene to the desired size by directly manipulating the colored Media Boxes within the scene outline. Students can easily learn this technique of one-to-one mapping of storyboard icon to media element. The direct manipulation paradigm for human-computer interface design influenced its conception.⁶

Users can also move media elements from scene to scene by dragging the Media Boxes from one scene to another in the scene outline. Users can play active elements such as sounds, videos, and animations by clicking on the media-distinctive "play" icon represented on the upper left corner of the element (Figure 4). Users can edit any media element by changing the style of its window from a display window to a fuller editing window, exposing the (limited) editing tools available for each type of media. For example, with text, users can control the font type and size. With sound, they can control the number of times (and with what volume) a sound will play when clicked. Double-clicking on its screen display changes the window style of a

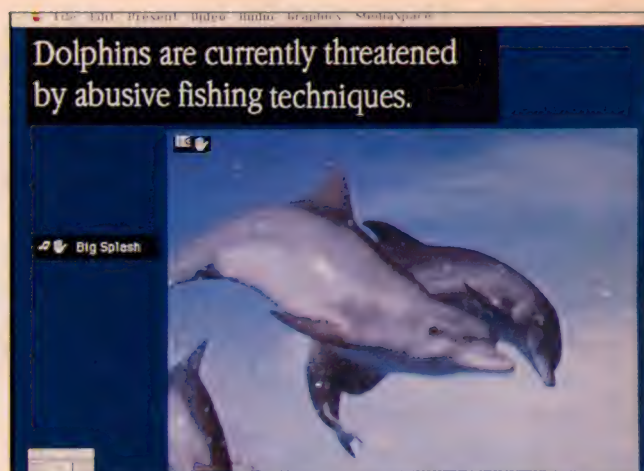


Figure 4. A display scene from a Multimedia Works document, with play icons for the sound and video elements of that scene.

media element from a display view to an edit view, and clicking on the button located in the upper right corner of that element's edit view window returns it to display view.

Three different processes present the composition. A standard method of using the Composer involves moving back and forth between scene design and viewing the media elements for that scene. Users design the scene on the storyboard and view the media elements for it on the display that appears behind it. To move from composing with the storyboard to presenting the scenes requires only a keystroke (down- or up-arrow key) to make the menu bar and storyboard disappear (or reappear). Right and left arrow keys are then used to navigate through scenes. The second presentation method employs the video-like controls on the media palette to step through or rewind the scene displays. The final method is to click the numbered buttons for each scene on the storyboard in succession.

For elements such as video clips, control palettes let you play these on an adjacent monitor (or in a video window on the Macintosh when using the Mass Microsystem's Colorspace Ili/FX videographics board). You can also use other commercial software programs in conjunction with Multimedia Works (see sidebar).

Video Light Table

Users can access the Video Light Table program under the video menu of the Multimedia Works Composer. We developed it because—unlike text, graphics, and sound—there were virtually no easy-to-use video editors available for the Macintosh environment. (Recently, Macromind announced Media Maker, which has some of these capabilities, and Digital F/X announced Video F/X, which combines video hardware and software for middle-end professional editing. The more expensive Avid system includes extensive use of image compression software for digital video storage.) We first developed the Video Light Table for use with the Pioneer 4200 videodisc

player. It also has drivers for Sony control-L protocol to allow for the common use of logging and editing 8 mm and VHS videotapes.

The user interface presents a log of the video as a spreadsheet of cels, as on a conventional light table (see Figure 5). In each cel, a digitized frame serves as an icon for that segment of videotape and includes start and stop times for the video segment. Users control player and recorder control panels and buttons for video digitization and assembly edits from a floating palette. Users can also sequence video segments (for either playback or recording) by spatially sequencing the video cels in the order they wish to view or record. To rearrange the temporal order of the segments, the user simply grabs the cels with the mouse and moves them to the desired location. The Video Light Table window displays the change.

Participatory design in learning

In developing a learnable set of multimedia tools, we began our effort with the following foundations:

- We focused on learner-centered environment design, which is emerging as a critical need from learning research in the cognitive and social sciences.⁷
- We focused on an iterative, participatory design process for our tools and materials. In this process, we considered

ethnographic observations of software use for real activities and user involvement in design as critical for creating a fit between tool and tasks.^{8,9}

Through their participation, students had a substantial influence on software design. They pressed for integrating the Media Space and Composer software programs, which we initially conceived of as separate tools. They continually suggested simplifications of the sequence of software operations that they had to go through to create media or design a composition. They led us to eliminate many functionalities we thought might be good ideas (such as special effects and dissolves between scenes); they did not perceive these functionalities as "core" to their design and production tasks.

Phase 1: Multimedia Works boys' club

Four different phases made up our Multimedia Works learning research. In May 1989 we began the first phase. With a boys' club of about eight adolescents in a local middle school from an underserved community, we started with preliminary designs in a rough Supercard prototype. The club met two afternoons each week for two hours until December 1989. Our objectives were to establish workable activities that would sustain the boys' interests, give us good feedback on the learnability and appropriateness of the tool's design for their needs, and suggest learning topics and activities.

Multimedia Works Hardware and Software Specifications

One media-creation/presentation station supports three composition stations. We recommend two media-creation/presentation stations together with six composition stations for a classroom's lab. Additional Multimedia Works stations consisting of a Macintosh II with 8 Mbytes of RAM, an 80-Mbyte hard disk drive and a 13-inch color monitor can share the audiovisual hardware.

Other requirements for the creation of Multimedia Works presentations include videotape and laser discs as media sources, blank media (diskettes, videotapes, and so forth) and electrical, video, and audio cabling plus the VPC cables to access the Control S and Control L ports on a VCR from a Macintosh computer.

Multimedia Works Station Hardware:

1. Macintosh II with 8-Mbyte RAM and 80-Mbyte hard disk drive
2. 13-inch color monitor
3. 13-inch multiscan monitor (for example, the Sony CPD 1302, if videographics board is not used)
4. 45-Mbyte cartridge hard drive (recommended for receiving upgrades and for exchanging digital compositions), plus three to four blank cartridges
5. Colorspace Graphic Overlay Board (Mass Microsystems)

6. Colorspace FX Video Board (Mass Microsystems)
7. Laserdisc player level III (for example, the Pioneer 4200)
8. Sony 8 mm VCR with 5-pin control port for video input (such as EV-S900)
9. Self-powered speakers (or TV monitor with speakers)
10. Copy stand and color video camera for digitizing images
11. Lavalier or Super-Directional microphone for narration
12. Video cassette recorder for recording output of presentations (such as the Sony SLV-757)
13. 8 mm video camera for recording video input

Software for each station (in addition to Multimedia Works software):

1. Macintosh Operating System 6.07
2. Supercard version 1.5 (Silicon Beach)
3. Smart Scrap Desk Accessory (Solutions International)
4. Soundedit and Macrecorder (Farallon) for sound digitizing
5. Director (Macromind) for creating animations
6. Colorspace Init & DA's: Driver, Desktop Video, NTSC Digitizer

Students developed multimedia compositions on their topic and presented them to other teams of students to explain, tell a story, or persuade. They spent about four to six weeks collaboratively researching and preparing presentations on a variety of topics in environmental science. At the conclusion of each cycle, they presented their work to their peers and recorded their presentations on videotape to take home or to school. Through this approach, students who created the compositions learned about both the subject matter and how to effectively communicate using written text integrated with other media such as graphics and video.

While you can use the Multimedia Works tools for virtually any topic, students' work examined issues in environmental science, including science, technology, and society issues like global warming, toxic wastes, and species depletion. Small groups of students researched and collected various media including newspaper articles, magazine illustrations, television documentaries, and their own video and audiotapes of field trips and interviews. With some guidance, students learned to employ critical thinking skills while conducting their own discussions. They learned to analyze the media they had collected, focus on their chosen topic, and select and organize the media to communicate their ideas. The computer tools then helped in composing texts, graphics, video, and sounds into a presentation.

Phase 2: Multimedia Works girls' club

One major change occurred when we began to work with a dozen middle school girls in a similarly organized club. We did more group work in planning the compositions; we spent more time researching, defining, and designing the media used in them. As with the boys' club, we focused on critical and goal-directed media use. But we had a new objective: to have students work on the use of argumentation and persuasion within the composition. We highlighted the importance of developing arguments that present evidence both for and against one's position. We also stressed the importance of weighing these considerations on behalf of the argument presented.¹⁰

The full group worked to brainstorm, choose composition topics, and divide up research work to define relevant media for potential use in their compositions. Refinements of design work and decisions were made in groups of two to three students.

In their first group work, the students composed a multimedia composition arguing against the planned construction of office towers in their downtown area. The presentation included digitized maps, graffiti artwork, traffic noises, newspaper sketches of the proposed structure, notes from interviews with neighbors on their perceptions of the development project, a video clip of an interview with a local realtor whose office would be displaced by the construction, and a videotaped speech on the group's beliefs, co-authored by the group but presented by one student. The students went on to develop compositions on topics such as the plight of various endangered species and their favorite popular music groups.



Figure 5. The Video Light Table Interface for video clip selection and assembly editing.

A fundamental improvement in the girls' club was the regular use of "design crits." Peers and researchers viewed an early draft of a multimedia composition and then offered constructive criticism of it. Specialists and experts on the topic the composition covered also joined in on some of these critiquing sessions.

Phase 3: IRL summer work

Enrique Godreau and Michael Chertok of the Multimedia Works Project ran an intensive research activity in July and August 1990. The Stanford Youth Employment Program provided funding for five students who had just graduated from the same middle school from which our club students had come. They worked for 20 hours a week at the Institute for Research on Learning on a number of Multimedia Works activities:

1. They created Multimedia Works compositions on how to use Multimedia Works, planned for use as training materials at the middle school in fall 1990.
2. They designed a report to the National Science Foundation using Multimedia Works. The report provided multimedia rationale and documented the research perspectives, software, and classroom-based research findings for a project on learning physics.
3. They created some open-choice documents, which included an impressive analysis of recent media events, such as flag-burning and the banning of rap-music lyrics, as tests of First Amendment rights.

During the same period, a number of teachers from the students' former school also came to IRL. They came to learn to work with the Multimedia Works software and rethink their media use in the classroom in terms of what the technology makes available (an effort supported by a Carnegie Corporation grant). This studio environment, in which students and

teachers worked side by side (without the traditional authority relations of the school context), provided a rich environment for mutual learning in their respective research and design efforts.¹⁰

Phase 4: Into the local middle school

One of our goals was to have the Multimedia Works environment migrate from an after-school club to regular use within middle school classrooms. But while the school had some low-end Macintosh computers, the higher-end Macintosh II and

Students who created the compositions learned about both the subject matter and how to effectively communicate using written text with graphics and video.

multimedia add-ons our software required were not available. Godreau and Chertok at IRL worked collaboratively with a development specialist from the school district on an Apple Computer Crossroads Grant application that would enable the school to support the transformation of its classrooms into multimedia studios that students and teachers could use as an integral part of their learning and teaching. The school serves more than 400 seventh and eighth grade students. This Apple Computer grant program is a competitive one, and schools must provide substantial evidence that they are at a "crossroads" in how they approach educational processes. In the case of our partner middle school, they were in the midst of a transition into "thematic instruction," an approach that combines formally distinct disciplines such as history, English, and social studies (or mathematics and science) through such activities as team teaching, project work, and longer class periods. The grant application succeeded, and in fall 1990, Multimedia Works began in earnest at the middle school.

Teacher training in summer 1990 and fall 1990 for the school-based Multimedia Works Project consisted of three components:

1. basic computer skills and general operation of Macintosh computer equipment;
2. multimedia production skills, including the use of specialized software and the operation of video and audio equipment; and
3. integration of multimedia production with the classroom curriculum.

The supervised practice in multimedia production was the most extensive portion of the training and consisted of helping pairs of teachers create two multimedia compositions to present in their own classrooms. Training included an exhibition of sam-

ple multimedia presentations, practice in the use of the software, instruction and practice in the use of audio and video equipment with particular emphasis on use of the video camera, supervised (and unsupervised) composition time, and group critiques of compositions in progress.

Working with multimedia tools

Since regular student multimedia activities began last fall, students have been filling the lab before and after school, at lunch, and between periods. I can briefly describe what the students have been doing by characterizing their processes of working with the multimedia tools:

1. Student teams are organized and assigned a topic. They discuss their topic and begin research using traditional resources such as newspapers, magazines, and encyclopedias. They also browse through the collection of images, text articles, sounds, and graphics that the teacher has assembled in the Media Space database.

2. Students take field trips, visit museums, and do homework assignments to gather media for their composition. In a project on endangered species, for example, students might videotape scenes at the zoo, record an interview with an environmental scientist working on these problems, use their VCR at home for recording a special news report on related topics, or find pictures of animals in books or a museum directory.

3. Student teams examine their collection of media, logging the contents of video and audio tapes and assembling the media they want to use to present their topic. An important part of this process is analyzing the media they have collected, that is, determining what a particular quote, picture, or video segment really conveys. The teams then begin to design and refine their media to plan an effective presentation.

4. Members of each team assume roles needed for the production of their presentation. Some of the primary roles include researchers, directors, storyboard planners, writers, graphic artists, designers, video producers, and audio producers. As a team, students plan the structure of their presentation using a storyboard.

5. The teams bring their media to the Multimedia Works composition station and begin to assemble their presentation. In the process of composition, the students continue to analyze the media they have collected, gather additional media, and refine the writing and organization of their presentation.

6. Once a first draft of their composition is ready, the student team critiques their own work together with their teacher and other participants who have special knowledge of the topics involved. Students create a list of what they want to do differently. The teams continue to revise their compositions.

7. Students share their multimedia compositions with other students in the class and record their revised presentation on videotape. Students answer questions and lead discussions regarding their work.

8. Students can present their work in other classes, in other

schools, or in a community forum. The videotaped presentations can be kept in a library collection of student-produced learning materials.

Some principal challenges

Certain issues are central to the successful use of Multimedia Works. These issues tend to lie at the intersection of technology design and user support for learning activities, and we are working on these challenges now:

1. At first, students often have difficulty comprehending the spatial representation of temporal media and distinguishing “plans” for compositions (that is, the live storyboard) from the compositions themselves. This would improve if a time-based controller for media displays within the Composer scenes on the Macintosh were possible. Currently, audio, video, and animation displays make synchronization a technical barrier.

2. While “closed” databases pose severe limitations for learning (students have to repurpose video, sounds, and so forth in their compositions), an approach that empowers students to extend such databases is not without its own problems. Moving toward multimedia-database extensibility by learners includes

- A difficulty in coming to “see” other media segments as data for use in a multimedia composition (that is, to conceive of part-whole relations). Students’ experiences of passively viewing TV and listening to music clashes with multimedia, where students can “capture” some segment of what they see or hear for subsequent use in a composition. The capacity for learners to use devices to capture live media such as audio and video is quite novel. These media distinguish multimedia compositions from writing and spur students’ imaginations in ways uncommon in text-centered composition.

- Attribution of media. New practices, somewhat akin to library scholarship, need to become a standard part of media use ethics by students. Otherwise, copyright issues will become a problem. Students do not always take the time to make references and attributions to the media they find or to cite other students’ media contributions to their work.

3. Although small-group learning activities are highly recommended in classroom learning research, getting our club participants to make reliable commitments and follow through to doing subparts of a larger multimedia development project proved difficult. The Multimedia Works software environment currently does not explicitly support project planning and monitoring activities. Often we need to request volunteers to do some component task of the larger project. Even with external guides to the part-task relation to the whole (such as a large white board that specifies a “storyboard” for the plan), the movement of students into subtask groups and the monitoring of progress toward these can be very hard to sustain. Part of this problem might have resulted from the club nature of the research.

4. Multimedia might fundamentally change the nature of

learning by creating rich conversational artifacts for discussion and presentation. To use the full expressive powers of different media, old molds of teaching and learning must be broken. One risk we fear is that multimedia technologies will be seen as sophisticated word processors. Rather than changing learning practices, they will fit into current systems of teaching, testing, and evaluation, thus perpetuating educational practices now under national critique.

5. Network support for collaborative activity in designing and critiquing Multimedia Works documents is desirable, but

An important part of multimedia composing is analyzing the media you have collected, that is, determining what a particular quote, picture, or video segment really conveys.

such support lies beyond the current budgetary possibilities of both the school and the project’s developers.

Implications

The research contributions of the Multimedia Works Project reside in the timely investigation of the educational opportunities and cognitive problems of multimedia. Our results indicate that we can easily make multimedia composition accessible to middle school students. The tools and activities promote rethinking cross-curriculum relationships and what students do to learn. We hope the “plan for success” project philosophy, exemplified in its formative research and participatory design methods, will result in a replicable model for student collaboration and small-group learning with multimedia.

Through the successes of this project, we hope to persuade researchers to develop tools for students to use for creating multimedia compositions. Students quickly learn to create such compositions and find the process engaging and motivating. We also hope to see educators using such learner-centered multimedia in the classrooms of the future.

We set out to support the choreographic sequencing of multimedia document elements in space and time. While the Multimedia Works interface provided clear support for these tasks, more time-precise displays of multimedia document elements pushed the current Macintosh platform’s capabilities. Students are demanding in this respect—they are accustomed to MTV and commercial film timing and rhythm and want their displays to reproduce these real-time effects. If we begin to think of the child’s desktop as the frontier of advanced computing rather than a trickle-down from military and industrial technologies, then the students might get their wish. □

BEST SELLING BOOKS

FROM



**IEEE Computer
Society Press**

VISUALIZATION AND IMAGING

Visualization in Scientific Computing

**- edited by Gregory M. Nielson, Bruce Shriver,
and Lawrence Rosenblum**

This tutorial covers the techniques by which scientists and engineers extract knowledge and form their results into complex and detailed simulations and computations. This book examines the techniques of scientific visualization derived from many areas including computer graphics, computer vision, signal and image processing, computer-aided design, and numerical analysis
304 pp. August 1990. Hardbound. Heavily Illustrated with 134 full-color photographs. ISBN 0-8186-8979-X. Catalog # 1979 List \$85.00 Member \$68.00*

Digital Image Processing and Analysis

- by Rama Chellappa and Alexander Sawchuk

VOLUME 1: DIGITAL IMAGE PROCESSING

VOLUME 2: DIGITAL IMAGE ANALYSIS

Volume 1 deals with image transforms and models, image data compression, image enhancement, reconstruction, and restoration. Volume 2 explores extraction of lines, curves, and regions in images, classification of objects, texture analysis, image analysis, scene analysis, and computer vision.

Vol 1 - 736 pp. 1985. Softbound. ISBN 0-8186-0665-7. Catalog # 665 List \$66.00 Member \$36.00*

Vol 2 - 680 pp. 1985. Softbound. ISBN 0-8186-0666-5. Catalog # 666 List \$66.00 Member \$36.00*

***All prices are 20% higher outside the U.S. and Canada**

MAIL YOUR ORDERS TO:

IEEE COMPUTER SOCIETY PRESS
10662 LOS VAQUEROS CIRCLE
P.O. BOX 3014
LOS ALAMITOS, CA 90720-1264

Or Call Toll-Free
1-800-CS-BOOKS

Or in California
714-821-8380

Or FAX 714-821-4010

(ADD \$5.00 PER BOOK TO COVER HANDLING CHARGES.)

Acknowledgments

The project and software were formerly known as "MediaWorks." Roy Pea created the project, and with Ed Boyle provided basic interface design for the Multimedia Works environment. Ed Boyle has been the principal programmer, with significant contributions from Enrique Godreau (Xerox Social Service Fellow) and Neils Veirum (Aalborg University, Denmark). Christina Allen, Chertok, and Godreau were instrumental in planning and conducting club activities. Other collaborators in Media Works research and development include Kathy Adams, Nancy Allen, Amitabh Ashesh, Isaac Jackson, and Jerome Shaw. Apple Computer's ACOT Project has supported the development of the Media Works prototype, and hardware for the project was provided by Apple Computer (External Research) and Farallon Computing, with some equipment support from Xerox Corporation. Partial support of the Science Media Works Club came from a contribution of the Carnegie Consortium Toward the Scientific Practice of Science Education, based at IRL and funded by the Carnegie Corporation. Student transportation was supported by the Hewlett Foundation. We especially thank our club participants for their participatory design efforts. Our colleagues at IRL were also helpful. This project began in February 1989. Presentations of some of this work have been made to conferences of the American Educational Research Association, CHI-90 invited poster, Computer-Using Educators (CUE), Mac World 90, Siggraph Hypermedia Exhibit, and the Social Science Research Council Workshop on Social Aspects of Computing.

References

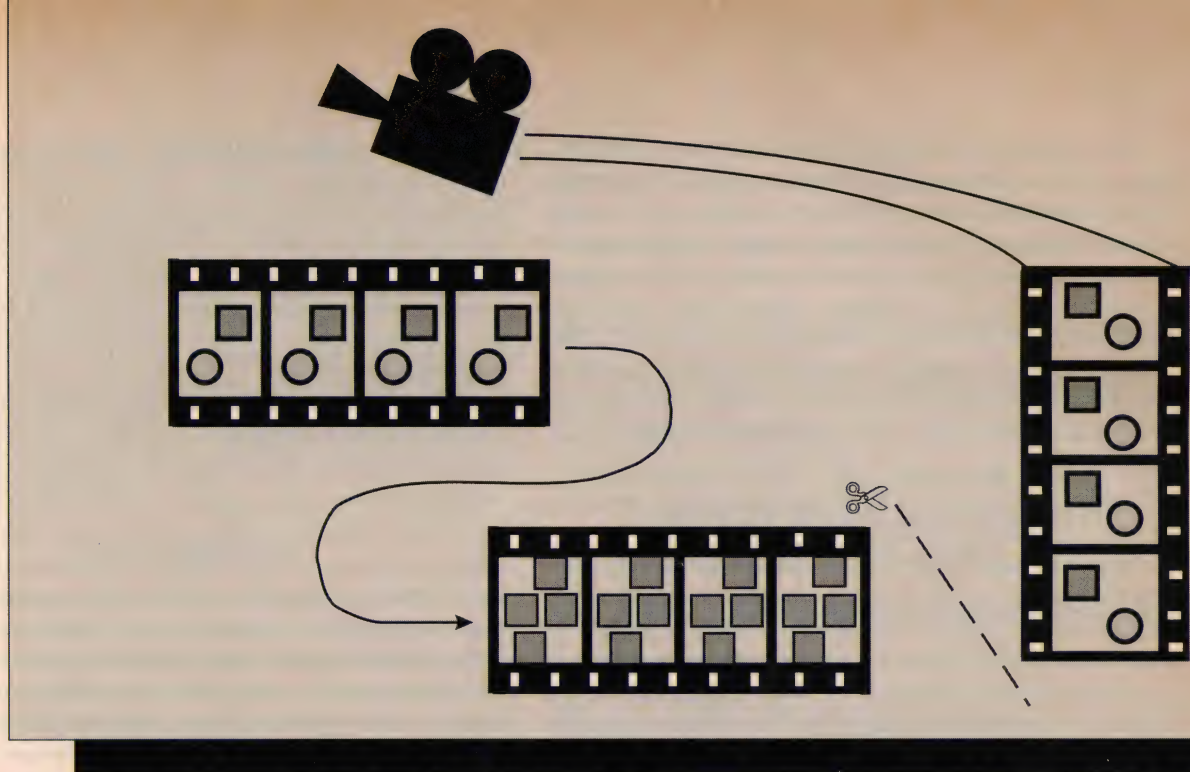
1. S.M. Eisenstein, *Film Sense*, Harcourt Brace Jovanovich, New York, 1947.
2. R.D. Pea, "Human-Machine Symbiosis: Exploring Hypermedia as New Cognitive and Cultural Technologies," to be published in *Cognition, Computers and Epistemology*, K. Plunkett and S.F. Larsen, eds., Lawrence Erlbaum Associates, Hillsdale, NJ.
3. R.D. Pea, and D.M. Kurland, "Cognitive Technologies for Writing Development," in *Review of Research in Education*, E. Rothkopf, ed., 1987, pp. 71-120.
4. R.D. Pea, E. Boyle, and R. de Vogel, "Design Spaces for Multimedia Composing Tools," in *Designing for Learning*, B. Bowen, ed., Apple Computer Press, Cupertino, Calif., 1990.
5. M. Mills, and R.D. Pea, "Mind and Media in Dialog," to be published in *Full Spectrum Learning*, K. Hooper and S. Ambron, eds., Apple Press, Cupertino, Calif.
6. J. Hollan, E. Hutchins, and D. Norman, "Direct Manipulation Interfaces," in *User-Centered System Design*, D. Norman and S. Draper, eds., Lawrence Erlbaum Associates, Hillsdale, NJ, 1987.
7. R.D. Pea and E. Soloway, "Mechanisms for Facilitating a Vital and Dynamic Education System: Fundamental Roles for Education Science and Technology," Rept. to the Office of Technology Assessment (National Tech. Information Service Rept. PB 88-194 634/AS), 1987.
8. C.L. Allen, "Reciprocal Evolution: A Design and Research Strategy for the Development and Integration of Basic Research, Design, and Studies of Work Practices," to appear in *Participatory Design*, D. Schuler, ed., Lawrence Erlbaum Associates, Hillsdale, N.J.
9. P. Ehn, *Work-oriented Design of Computer Artifacts*, Almquist and Wiksell Int'l, Stockholm, Sweden, 1988.
10. C.L. Allen, "Multimedia Learning Environments Designed with Organizing Principles from Non-School Settings," to be published in *NATO Advanced Research Workshop on Computer-based Learning Environments and Problem-Solving*, E. De Corte, et al., eds., Springer-Verlag, New York.



Roy D. Pea is a senior research scientist at the Institute for Research on Learning and a consulting professor at Stanford University. He is a cognitive scientist with special interests in integrating research and the design of effective learning environments for science, programming, and multimedia computing. In September 1991, he will become John Evans Professor of Education and the Learning Sciences at Northwestern University.

Pea received his D.Phil. in developmental psychology at Oxford University. He is a member of the American Educational Research Association, the Cognitive Science Society, the American Psychological Association, and SIG CHI.

Readers may contact Pea at Institute for Research on Learning, 2550 Hanover St., Palo Alto, CA 94304.



Cinematic Primitives for Multimedia

Glorianna Davenport, Thomas Aguiere Smith, and Natalio Pincever
MIT Media Laboratory

*Shots, sounds, and
sequences advance the
storytelling in film.
Stratification helps us
use these elements in
multimedia.*

Throughout its history, cinema has challenged filmmakers to experiment with new storytelling modes. Interactive multimedia proposes that the participant viewer can affect selection and sequencing of cinematic story elements. Our challenge is to develop robust frameworks for representing story elements to the machine such that they can be retrieved in multiple contexts. We propose that content can be represented in layers, or *strata*. This model for layered information will allow programs to take advantage of the relation between cinematic sequences and the world they represent.

Interactive storytelling

Interactivity between participant users and machine representations is an integral element in today's multimedia computer environments. Thus, interactive multimedia is a user-directed form of storytelling. Whether we talk about hypermedia publications, digital video editing systems, real-time collaborative exchanges, or personalized access to large on-line video servers, we want to offer participants the most natural possible access to information. Computer programs coordinate the interaction by mediating between user input and segments of multimedia data. User navigation through multimedia systems and successful computer mediation of the data in such systems depends on an effective representation of the content.

Today, after 12 years of developing laboratory prototypes, companies have begun to distribute commercial multimedia applications. ABC News Interactive in New York, WGBH in Boston, Voyager in Santa Monica, Synapse Technologies in Los Angeles, and Time-Warner New Media in Los Angeles number among current publishers of interactive video programs. Museums increasingly commission interactive media pieces, including the Exploratorium in San Francisco, the Baseball Hall of Fame in Cooperstown, New Jersey, the Computer Museum in Boston, and the Smithsonian in Washington D.C. While many of these applications constrain user navigation by adherence to the author's preprogrammed cues, others are more open ended and encourage browsing.

In browsing, users' personal goals and intentions may collide with the system's ability to foster coherent meaning; viewer comprehension tends to break down, therefore, when participants begin browsing. While this breakdown is often blamed on weaknesses in the navigational tools, it more frequently reflects the lack of a semantic representation that can support the limited look-ahead functions required to build meaningful interactions with the user.

Increasingly, we will need to create multimedia systems that maintain a conversational mode of interaction with users by generating and tracking story frameworks. In the case of on-line video servers as well as home-movie editing assistants, the machine must respond to the user by selecting the "best" shots, sounds, and text chunks, then orchestrating or sequencing them to emphasize a particular story. The story content should reflect the user's background and intent.

Factors that affect multimedia's conversational success include the quantity of video and audio available in a system (largely solved once low bandwidth digital video becomes ubiquitous), machine-readable representation of the program content, articulation of complex narrative models, and natural language understanding. Significant advances in all of these areas continue to increase the power of multimedia information and simulation systems.¹⁻³

Developments in cinematic storytelling must join technological advances in multimedia if we are to produce interesting machine-mediated experiences. The history of cinema offers a sensual and provocative record of the evolution of an expressive form. A rapid overview of this history in light of technical innovation allows us to examine the traditional levels of granularity—that is, dividing a stream of information into meaningful chunks—that have been used for the moving image.

The problem might be quite obvious but bears restating: For a machine to select and sequence video footage, the content of footage must be represented in machine-readable form. Today, no machine-readable information, with the exception of SMPTE (Society of Motion Picture and Television Engineers) time code, accompanies motion picture images as generated. For the purposes of archiving and editing, logs are sometimes created after collecting the footage. Such logs are time consuming to construct, suffer from a range of information errors, and, most importantly, lack the rigorous methodology needed for computer representation. To realize our vision of conversational multimedia, we need to attach known descriptive attributes to specific shots and sounds as we record them.

Some useful terms

granularity

Granularity refers to the descriptive "coarseness" of a meaningful unit of multimedia information. For image processing, the descriptive granularity has to be fine enough to address specific objects in the frame. When editing a movie, the granularity must be coarser to encompass thoughts, actions, and intentions.

chunks/segments

By chunks we mean segments of arbitrary length, usually with some semantic significance. The coarseness of these chunks determines the granularity of the system.

log/logging

Logging is the process through which shots in a piece of footage are cataloged, usually by in- and out-points, length, and a short description. Such a catalog is called a shot log, or log.

interactive multimedia

We like to characterize interactive multimedia as a user-directed form of conversational storytelling. In

multithreaded narratives, the user affects the "plot" of the story through choices along the story space.

shot

This is the smallest addressable unit having the finest level of descriptive granularity. It consists of one or more frames generated and recorded contiguously and representing a continuous action in time and space. The single frame is the smallest meaningful entity.

sequence

A collection of shots no longer perceived as a set of individual shots, forming a natural unit. This unity is the result of continuity in various planes: temporal, spatial, perceptual, and so forth.

An extensive glossary of terms for film and media can be found in James Monaco's *How to Read a Film: The Art, Technology, Language, History, and Theory of Film and Media*, Oxford University Press, New York, 1977.

Cinema and storytelling style

What makes a movie complete? What distinguishes a scene from a sequence? How do sound and picture complement each other to produce a seamless structure? With each edited production, narrative or documentary, filmmakers gain new understanding about how the pieces of cinema—shots and sounds—go together and how their progression affects the audience. As technology evolved, filmmakers invented new aspects of cinematic language. In this process, they reshaped the relationship between specific cinematic elements, the final movie, and the audience.

The silent era

In the beginning, motion pictures were shown without sound. Eisenstein's theory of montage emerged during this time. In the montage approach to editing, the narrative results from a carefully structured linear ordering of shots. The meaning arises from the associations made in the viewer's mind while watching these visual juxtapositions.⁴ Sets of shots compose a visual sequence that frequently suggests a psychological state or thought progression. This storytelling method relies on metaphor more than on continuity of action, on associative context rather than on directed context.

Filmmakers of this era who needed to develop the action frequently used text cards to further the plot. They shot all images with prime lenses, since zoom lenses were not yet a staple of the setup. The idea of visual, spatial, and temporal continuity developed slowly.

Movies with sound

As theaters became outfitted for projection, presenters began to accompany movies with live music, most frequently piano, selected to induce particular reactions from the audience. These live accompaniments were usually free-form, not specifically rehearsed or synchronized to the action. As soon as it became technically possible to lock sound to the moving image for projection, sound became the ubiquitous partner of the motion picture.

The advent of sound heralded the death of montage and the rise of movies that depended on visual continuity and specific dialogue to further the action. While the audience perceived this track as synchronous with the picture, most release tracks for narrative film were (and still are) created in dubbing studios and foley pits during the postproduction process.

As the tools for production and presentation improved, so did the language of the moving image. Directors began to use perspective as a vehicle for manipulating audience reaction to story elements. The camera became coconspirator with the creative vision of the cinematographer, who demanded increasing flexibility. New lenses were invented, and the pressure was on to make the camera more mobile, as well as to discover new, faster film stocks. These advances in technology led to the construction of much more sophisticated sequences. Movie directors became fascinated with realism and almost overnight invented

conventions that allowed viewers to build cognitive maps of the story space. The development of higher film speeds had a profound effect on the evolution of new cinematic techniques.⁵

Throughout the evolution of cinema, directors, cinematogra-

Descriptions of content must be structured at the shot level to maximize the potential for computer-aided browsing and sequence assembly.

phers, and editors sought to discover ways in which cinematic language could both camouflage and create a perceptual awareness of time and space. Whether match cut or jump cut, the relationship between the last frame of one shot and the first frame of the next shot generated new theories of continuity and ideas for sequence construction.

Granularity of meaning

In motion picture parlance, syntax refers to how action is framed in individual shots and how those shots can be ordered into sequences. While the audience experiences linear movies as unified entities, filmmakers experience movies as the generation of individual shot and sound elements and the forging of relationships between these elements.⁶ At each step in the production process, key professionals—directors, cinematographers, editors—contribute their knowledge and vision to the design of shots, sounds, and sequences.

In multimedia applications, especially those that draw on centralized databases of visual information, we will need to adjust levels of granularity as we incorporate the same piece of footage into different contexts. For instance, if we want to learn about recent events in Poland, we might choose to view sequences introducing us to the country, its language, and its culture. Later, we might want to build a country report using information about the country with recent political news. Clearly, we need descriptions that present us with longer or shorter elements according to the detail our query requires. Descriptions of content must be structured at the shot level to maximize the potential for computer-aided browsing and sequence assembly.

A shot consists of one or more frames generated and recorded contiguously, representing a continuous action in time and space. In most instances, the beginning and end frames of a recording medium are used as an address for shots. We can compute the duration of a shot from these two numbers and a designated playback speed. Obviously, the original shot generated in the camera can be redefined many times during editing.

Frequently a passage of sound is recorded along with the shot (either on the same or on a separate medium). We refer to this as *synchronous sound*. Synchronous sound consists of semantic dialogue and ambient sound. If this synchronous sound track is not physically bound to the picture (as in film or digital modes, which use separate picture and sound servers), it must carry some identifying tag to guarantee that it will maintain its synchronous relationship with the picture despite adjustments during editing. If you want to be able to relocate the original shot or synchronous sound passage from an edited program, a trace

We urge you to contemplate the role that the environment serves in its several guises. The environment serves as the glue we can use to generate a description for all aspects of the shot.

identifier must be included as part of the descriptive data set.

We can describe a shot in four general ways, each linked to the others:

- Perspective: Whose point of view does the image represent?
- Camera and sound recorder: What is the camera or microphone doing relative to the scene? Or how is the camera or microphone mediating the environment around it?
- Content: What is happening in the environment as it is captured in the image and sound?
- Context: What additional meaning is represented in the shot based on its adjacency to other shots or to related world knowledge?

Figure 1 displays a map of the shot. This map indicates how much information we, as editor or viewer, must understand to make sense of the image before us. In this global view, we extend each component of the shot—perspective, camera, content, and context—to show the extraction of this aspect. We urge you to contemplate the role that the environment serves in its several guises. The environment serves as the glue we can use to generate a description for all aspects of the shot. It is a cross-reference point for our structured model. By rotating a three-dimensional layout of the model, we should be able to superimpose the several instances of environment and move between the levels of data corresponding to each of the aspects.

Collectors of content: Camera

While Figure 1 lays a framework for understanding the complexity of the shot, it does not clearly reflect its temporal nature.

In reality, all of these elements are subject to change in time—as the shot is generated. This makes the idea of describing the image appear as a gargantuan task. However, if we turn to Ricky Leacock's description of a movie as "the filmmaker's perception of what takes place in the presence of a camera when the camera is turned on," we are reminded that what is captured on film involves the relationship between what happens in a given environment and the physical presence of a camera as it records over time in that environment. We can exploit the temporal nature of this process to track changes in state of both the camera and the images that the camera records over time and to retrieve this description with variable granularity.

This argument considers the camera a physical object, situated in space and time. The camera records images through a single lens to a discrete image plane. When we begin a shot, we know the date and exact time. We know the camera location, lens, and direction of view (more or less exactly, depending on our level of instrumentation). Moreover, the camera person knows the general intention of the filming and can identify, with some precision, the significant people and objects in proximity to the camera. We refer to this surround as the ambient reality of the camera (see Figure 2).

The physical act of the camera motion, of light passing through the lens to the image plane, and of a specific number of frames being generated over time, imply that the ambient reality established in whatever detail at the head of a shot will, if properly encoded, be constant throughout the shot except as changes of state occur. Changes of state happen either to the recording equipment or to the objects inhabiting the particular environment in which we are recording. Therefore, we can represent information about the ambient reality to the computer over time in as much detail as we wish.

Stratification defines a model of layered information we use to describe what is happening in a shot.⁷ We can attach a range of attributes including frame delimiters, who, what, when, where, why, and so forth over the temporal span of the shot. We can attach more generalized descriptions to the scene level or program, and these will be inherited by the shot. Likewise, the program as a whole can inherit information specific to a shot.

To keep better track of state changes, we are currently constructing a graphical logging device that will allow us to annotate the layered information related to video footage with greater ease. The variable granularity will allow users or programs to query the database and to identify shots and sequences in a temporally appropriate scale. The model we suggest links inheritance to granularity.

Collectors of content: Microphone

As with the camera, the microphone or sound recorder mediates the environment in which it is situated. Ambient sound offers us clues about the physical and geographic aspects of place. When parsed, the original synchronous sound track for

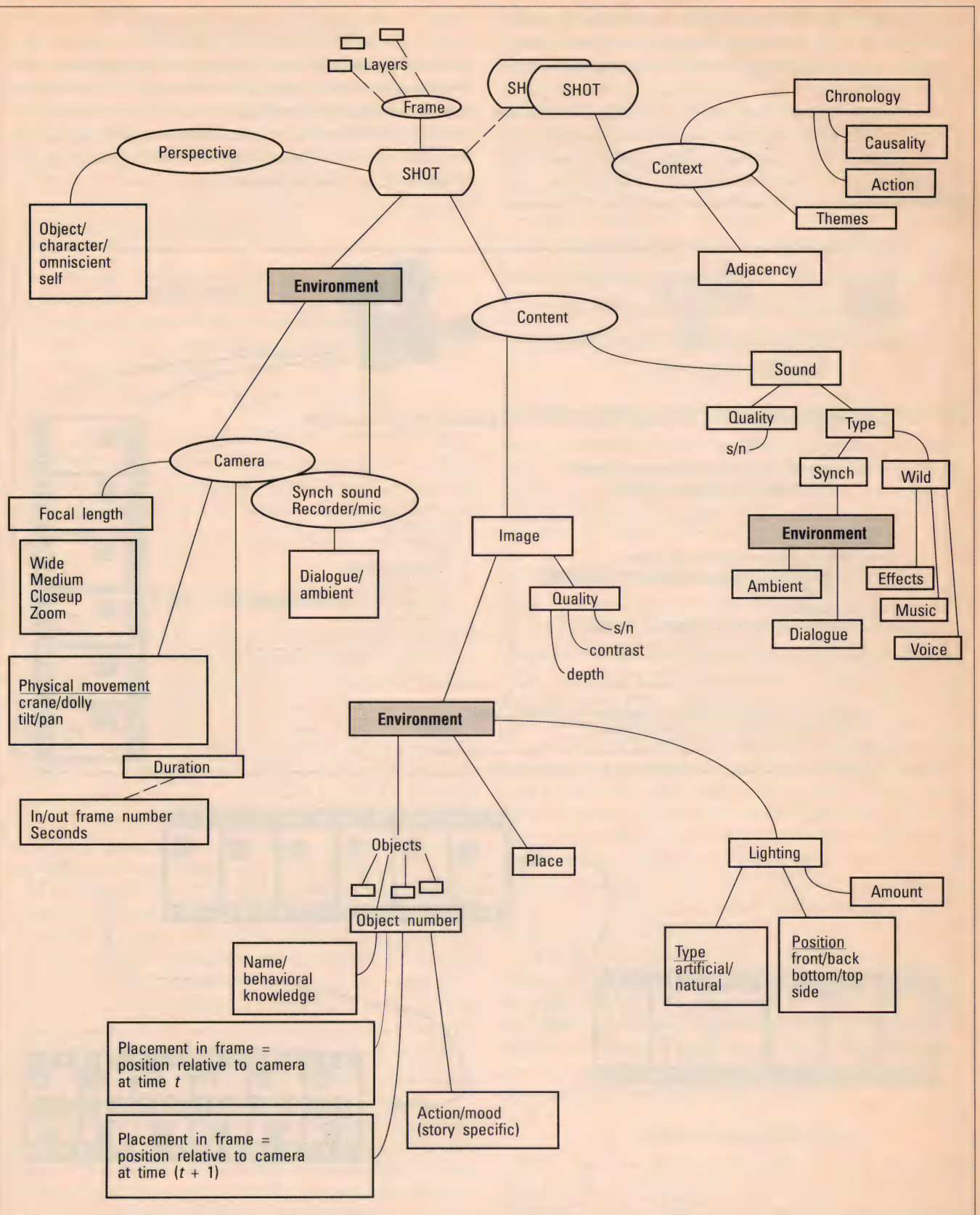


Figure 1. The shot: toward a structured model. We break down information contained in a shot in terms of perspective, camera and sound, content, and context. In each case our understanding of the relation of the recorded information to the environment helps us articulate meaning.

any picture can help us segment the video/audio stream into shots or help us determine significant action points. Notwithstanding, sound poses some difficult challenges to our representation methodology.

For instance, in interviews sound is usually the principal source of meaning. To represent the content of the interview, which is frequently captured as a single shot, the text of the interview needs to be broken down into conceptual units. The

length of these units can vary enormously. While text transcripts of interviews and of all dialogue offer certain value to multimedia manipulations, neither speech recognition nor natural language understanding are sufficiently developed to help in constructing a representation from real-world dialogues. Inevitably, we will need to encode a summarized conceptual representation of dialogue tracks as part of our stratified representation of movie content.

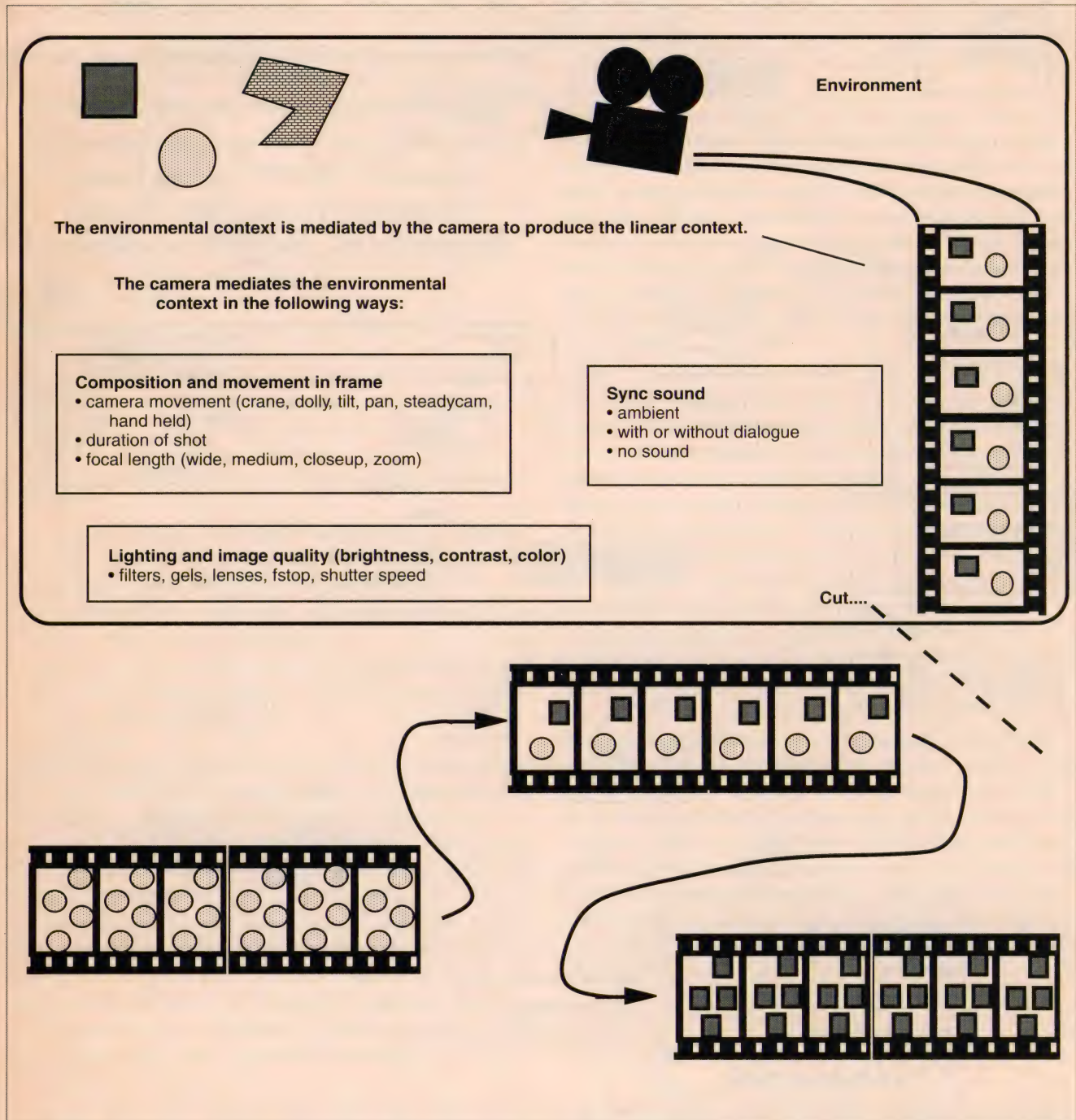


Figure 2. The situated camera. In editing, you can create context (chronology, causality, action) and develop themes by ordering image units into sequences; employing optical effects such as fades, dissolves, and wipes; using sound such as ambient, live, voice, and special effects; and adding text and break points.

The data camera

Some readers might already know that the cost of data entry for this stratified description will quickly become overwhelming if we rely only on after-the-fact human annotation. Because our representation method relies on the idea of the situated camera, we are currently considering how to build a camera to collect some of our descriptive information during the shooting process. We call this camera a *data camera*. Such a camera will record multiple tracks of data, including

- time code (SMPTE is already generated in professional electronic recording systems);
- camera position (by incorporating a global positioning system, as well as dynamic tracking of focal length and focus setting recorders); and
- voice annotation of who-what-why information.

Such data tracks can easily interface to a stratified logging system. Information collected during shooting is more reliable than after-the-fact annotation and will equally serve both modes of linear editing and interactive multimedia browsing applications.

The sequence

While the shot and associated audio are primitive elements, the sequence represents a second order of cinematic structure. Sequences are created by placing one shot next to another. While each shot describes a part of an action, the sequence allows us to understand a whole action, event, or thought. Whether comprised of one or more shots, a sequence is distinguished by the fact that it is no longer perceived as a set of individual shots. The transition between shots may be further blurred by the existence of a mix track. We believe that both the machine and the individual editor can use our descriptive methodology to build meaningful, context-rich sequences.

In traditional cinema, the editor tries to create a seamless experience. As just mentioned, sound can be used to blur distinctions at the edges of shots, sequences, and scenes. As a final step in cinematic postproduction, after the editor has fine-tuned or "locked picture," a sound editor is frequently employed to "clean up the track," embellish the feeling of being there by inserting sound effects, and create audio transitions. In addition to sync sound, a sound editor can add a track providing explanatory commentary or music to set a mood, to make information more explicit, or to heighten the perception of an action.

The utility of sound effects on reel can clearly be extended into multimedia systems. Sound libraries, such as the "Sound Ideas" CD collection, are currently being marketed. As we develop models for sequence construction, we need to be able to link to sound libraries in order to build seamless sequences.

One of the goals of "conversational" interactive multimedia systems is to create coherent linear sequences on the fly, according to user preferences. As sequences are built, the context

inherent at the shot level is harnessed and made explicit. In both documentary and narrative movies, context is developed through manipulation of chronology, introduction of causality, or simply structuring a completed action. Interactive multimedia differs from traditional cinema in that participant viewers may need to choose between multiple paths, may need to review their course, and may wish to actively restructure a se-

Stratification is a model for annotation that approximates the way in which the editor builds an understanding of what happens in individual shots.

quence. Nevertheless, both the viewers' experience and associated learning proceeds in time and can be enhanced by the generation of seamless narratives. This requires dynamic look-ahead and scalable descriptions.

Using representative strata

While movie viewing is sequential, movie production is inherently nonlinear. To construct coherent sequences, the editor must understand what took place in the environment mediated by the camera. As part of this process the editor builds a cognitive map of the environment and keeps track of who, what, when, where, and why over the temporal duration of a shot.

Stratification is a model for annotation that approximates the way in which the editor builds an understanding of what happens in individual shots. Much of this information could be collected quite easily by a data camera at the time of shooting. Therefore, stratified annotation appears to have several advantages for multimedia environments. First, it can comfortably support granularity. Second, other programs can easily parse the descriptions. Third, the original annotation can be tracked through multiple representations.

In addition, stratification can help support computer-generated links between similar pieces of footage. Editors use linking to associate sequences when creating linear narratives. Stratified descriptions of content will therefore expedite on-the-fly creation of machine-generated, multithreaded parallel narratives that users can adjust through linking. Such narratives are vital for effective multimedia environments. □

Acknowledgments

Ideas in this article evolved over the past six years. They result both from building multimedia systems and from producing multimedia publications. While we have given a few directly relevant references, many other students and research affiliates of the Film/Video Group (recently renamed the Interactive Cinema Group) contributed to our exploration and understanding of cinematic primitives. A short list follows: Dan Applebaum, Hal Birkeland, Brian Bradley, Hans Peter Brøndmo, Amy Bruckman, Ben Rubin, and Martha Swetcoff. Rele-

SPECIAL DISCOUNT PACKAGE

FROM



IEEE Computer Society Press

GRAPHICS AND IMAGING

Visual Programming Environments: Paradigms and Systems

- by Ephraim P. Gliner

676 pp. ISBN 0-8186-8973-0. Catalog # 1973
List \$90.00 Member \$72.00

Visual Programming Environments: Applications and Issues

- by Ephraim P. Gliner

702 pp. ISBN 0-8186-8974-9. Catalog # 1974
List \$95.00 Member \$76.00

Computer Graphics: Image Synthesis

- by Ken Joy, Charles Grant, Nelson Max, and Lansing Hatsfield

380 pp. ISBN 0-8186-8854-8. Catalog # 854
List \$60.00 Member \$40.00

Computer Graphics Hardware: Image Generation & Display

- by H. Reghbatl & A. Y. C. Lee

384 pp. ISBN 0-8186-0753-X. Catalog # 753
List \$50.00 Member \$35.00

1st Conference On Visualization '90

520 pp. ISBN 0-8186-2083-8. Catalog # 2083
List \$98.00 Member \$49.00

◆ PACKAGE PRICE ◆

**Order # GIP — List Price \$235.00
Member Price \$190.00**

MAIL YOUR ORDERS TO:

**IEEE COMPUTER SOCIETY PRESS
10662 LOS VAQUEROS CIRCLE
P.O. BOX 3014
LOS ALAMITOS, CA 90720-1264**

or call toll-free
1-800-CS-BOOKS

or in California

714-821-8380

or FAX **714-821-4010**

(ADD \$5.00 PER BOOK OR \$15.00 PER PACKAGE TO COVER HANDLING CHARGES.)

vant research was conducted under grants from MIT's Project Athena and in collaboration with research affiliates from the Asahi Broadcasting Corporation, particularly Keishi Kandori, Kenzo Furukawa, Yoshio Obata, Toyokazu Yoshida, Kuniyoshi Chihara, and Hiroshi Ikeda. Thanks also to Ricky Leacock for ideas communicated in many conversations and lectures between 1978 and 1988.

References

1. S. Morgaine, *MIDS: A System for Describing Image Content for Multimedia Design*, MSVS thesis, MIT, 1989.
2. S.M. Stevens, "Intelligent Interactive Video Simulation of a Code Inspection," *CACM*, Vol. 32, No. 7, July 1989, pp. 832-843.
3. W. MacKay and G. Davenport, "Virtual Video Editing in Interactive Multimedia Applications," *CACM*, Vol. 32, No. 7, July 1989.
4. S.M. Eisenstein, *Film Form*, Harcourt Brace and Co., New York, 1949.
5. G. Toland, "How I Broke the Rules in Citizen Kane," in *Focus on Citizen Kane*, R. Gottesman, ed., Prentice-Hall, New Jersey, 1971, pp. 73-77.
6. K. Reisz and G. Millar, *The Technique of Film Editing*, Focal Press, Boston, 1978.
7. T.G. Aguiere Smith, "Stratification: Toward a Computer Representation of the Moving Image," a working paper from the Interactive Cinema Group, MIT Media Lab, 1991.



Glorianna Davenport is an assistant professor of media technology, first holder of the Asahi Broadcasting Corporation Career Development Chair, and Director of the Interactive Cinema Group, a research group at MIT's Media Lab. Formerly an independent documentary producer, Davenport has been involved in the development of computational multimedia since 1982. Her work includes interactive programs such as New Orleans Interactive and the Elastic Charles, as well as random access video viewing and editing tools.

Davenport received the Gyorgy Kepes Fellowship Prize for excellence in the arts at MIT in the fall of 1990. She holds an MA from Hunter College and a BA from Mt. Holyoke College.



Thomas G. Aguiere Smith is currently a masters degree candidate in the Interactive Cinema Group at the MIT Media Lab. His research interests include multiuser networked video databases, problems related to on-line video archiving, video indexing systems, and how video can be used as a research tool. His current research involves developing a computer representation of moving image content called "Stratification."

Smith received a BA degree with a concentration in anthropology from the University of California at Berkeley.



Natalio Pincever received his MS in media arts and sciences at the MIT Media Lab in 1991. His current research focuses on cinematic parsing of audio tracks for the Interactive Cinema Group. His research interests include digital audio processing, desktop audio, advanced human interfaces, multimedia communication, and hypermedia.

Pincever holds a BS degree in electrical engineering from the University of Florida. He is a member of the IEEE Computer Society.

The authors can be reached at the Interactive Cinema Group, E15-432, MIT Media Laboratory, 20 Ames Street, Cambridge, MA 02139.

IEEE Computer Graphics & Applications

MULTIMEDIA

The Demise of Multimedia

David R. Clark

University of London Audio-Visual Centre

*Forget "multimedia."
We can best serve the
future of interactive
electronic presentations
(IEPs) by developing
special-purpose delivery
engines rather than
general-purpose
computers.*

The naming of things is of prime importance. It provides a guide to the origin and context for the object so named and helps those who would use it to understand it. We must therefore delete the term "multimedia" from our language. The word is deeply unhelpful for three reasons: grammatical, in that anyone with an ear for language cannot use it as a singular noun; semantic, in that it has resisted adequate definition; and philosophical, in that it in no way expresses the intentions of those whose activity it purports to describe. We urgently need to find a word adequate to characterize intelligibly those activities now referred to as "multimedia."

For this article I use "interactive electronic presentation" (IEP) to describe a collusion of sounds and images elicited from a piece of (electromechanical) machinery by the user's persistent activity. Insofar as the sequence of events so elicited generally follows the intentions of an author, there usually exists an intention to convey. In this sense the device and its associated programming mediate a message, and we can safely describe it as an "interactive electronic medium": Conveyance of the message partly by pictures and sounds is a matter of convention, and drawing attention to this fact by including such phrases as "audiovisual" over-eggs the pudding. The message is mediated electronically, and the plethoric implication of the prefix "multi-" is unwarranted.

Mediated communication

A telephone conversation between two people is an electronically mediated, interactive, synchronous communication. This article addresses the construction and delivery of presentations that are asynchronous in the sense that the owner and creator of the message are not cotemporal with the receiver. In short, I will discuss books in a new form, and questions of storage and retrieval thus become essential.

Separating creation from delivery

In the case of a (conventional) book we readily accept that the processes of creation and consumption differ radically. The small convenient object that works on the subway and in the bath and is cheap enough to form part of a private collection is the product of a large and expensive printing press. Press input, in a specialized form unconnected with the eventual output, results from many iterative design cycles, again using devices unrelated to the final object. The same holds true for a film or television program, where the diverse resources of the studio and editing suite contract to the specialized resources of a single video player by the time the final product reaches the end user. This same disjunction will prevail in any new form of mediated communication—the engines for design and delivery will differ.

However elaborate, an interactive electronic presentation is closed and finite. It got that way as a result of a design process that began with a much more open and inchoate aspect. This implies that the resources required to sustain the mediation of the final object are precisely specifiable in a way that the resources for the creation of the object are not. The delivery engine must perform a set repertoire of tasks as efficiently as possible, therefore it can be optimized against a defined set of criteria. The creation engine will have a different set of constraints against which it must be optimized and will thus take a different form.

Unfortunately, something not bad for everything is not good for anything.

Optimizing against constraints has an important consequence: The resulting object, while well-suited for the designated task, becomes severely handicapped once the original constraints are violated. A good illustration of this is the attempt to exploit a technology designed to fool the ear for the entirely different purpose of fooling the eye. The constraint of delivering an uncorrupted digital datastream of sufficient density to reproduce the variation in air pressure associated with frequencies in the range of human hearing produced the compact disc digital audio (CD-DA) system.¹ Attempts to use that

storage and delivery system to deliver digital data at rates sufficient to provide moving images of arbitrary complexity on screens of an adequate size for the human eye are, at the very least, frustrating—the tool does not easily fit that purpose. Commercial imperatives can, however, force such an activity.

In an attempt to escape the constraints of specific optimization, machines today are designed to be as general purpose as possible. Unfortunately, something not bad for everything is not good for anything. Since it appears to be a general truth that successful development tends toward diversity and specificity,

Something becomes information only when set in a context informed by a purpose.

we can expect the next generation of engines for personal use to be tailor-made for particular tasks. For example, the best computer in my house is the one that controls the washing machine—it does very little, but what it does, it does well. These considerations imply that any delivery system for IEPs will be highly optimized, not a personal computer in today's sense of the word, because of the demands associated with the storage and manipulation of large numbers of images.

Image demands

A characteristic of the eye is its facility for detecting “edges.” The visual artifacts introduced by the sampling required to enter the digital domain often create more problems than the benefits conferred by digital representation. For a great many images the noisy imprecision of the analog video representation is more than adequate and less distracting than its conventional digital representation. Furthermore, the technology available to deliver analog images has a far greater narrative versatility, in that we can easily access and deliver arbitrary stills from some 50,000 per disc in about 1 second and play motion sequences forward and backward at a wide range of rates.

While great strides have been taken in the compression and decompression of digital data, I argue that the benefits of storing the data in digital form do not, in the majority of cases, outweigh the limitations on narrative versatility that this form necessarily imposes.

Framework for delivery

Ever since the pioneering work of Negroponte et al.² at the Architecture Machine Laboratory (now the Media Lab) at MIT, it has been clear that if we could bring moving images and sound under computer control, then interesting extensions to our ways of assimilating information would result. In fact, very little of what exists today in the way of ideas and concepts in this

field was not first explored in that group. Unfortunately, we need to shift the paradigm established by this work before we can see the full power of the concepts clearly. We must remove an essential misconception: the notion of information as autonomous and “browsable” for its content. Something becomes information only when set in a context and informed by a purpose.

To see this clearly, consider the four million or so photographs of London from the time of Fox-Talbot to the present held in the archives of the Museum of London. According to the curator, the images in this collection have just two things in common: they are all photographs of London, and they are never accessed for the reason for which they were taken. There is thus no single way to catalog them according to their content. Yet surely they constitute a rich source of “information”? Unfortunately, they only become so when someone strings them together in a narrative to convey a thought. Human thought gives the images meaning in that particular context. Those same images could quite possibly form the backbone of narratives totally unrelated in terms of their informing concepts (for example, the development of the London Docks or the efficacy of photographic fixing agents).

This illustration emphasizes the central role played by a book as deliverer of an argument. The notion of self-assembly of a truly personalized idea from a kit of parts is misleading: Either the kit only contains those elements that force a particular type of argument (all Lego structures are essentially the same, however varied they try to make the bits) or vital pieces needed for the development of an unforeseen argument are missing (since no finite container can contain all possible instances).

Once we abandon the quest for the management of arbitrary diversity, the real goal becomes clear. Just as with all previous inventions for asynchronous communication, this new one must be specifically adapted for storytelling. The IEP creates a new dimension, allowing users to reconstruct the author’s intended narrative in their own ways.

In film and video media, time never stops. The narrative flows on irrespective of the user’s state of mind, and the question of interactivity doesn’t arise. With the possibility that the narrative can become connected to the user’s personal time come previously unlooked-for opportunities. These do not lie so much in the ability of users to pursue their own interests but rather in the increased scope for storytelling sensitive to the differing preoccupations of would-be users.

The creation engine

All the difference in the world exists between devising an IEP and using it. One of the great unlearned lessons of this kind of electronic production—more than in any other medium—is the vital importance of detailed planning. As yet there exist few tools suitable for the modeling of interconnections between such disparate objects as pictures of the real world, diagrams, speech, music, computer graphics (whether animated or not), and written text.

The creation process is a multilayered exercise. Disparate resources must be available in browse mode as well as for inclusion in some subsection of a sequence of a larger module that can be run independently of its antecedents as a trial, then changed or incorporated into the growing whole without conflict. These tasks put a premium on good programming practice. (And nothing so degrades machine performance as good programming practice—all that self-commenting code with the correct use of local and global variables that leaves the stack as you would like to find it.)

These requirements of the creation process bring to the fore the functionality of what we might call the IEP constructor’s workshop. They begin by specifying 8 Mbytes of RAM; 500 Mbytes of (removable) disc storage; full multitasking; complete integration of computer-generated and external video, including frame-grabbing in color and the management of laser disc

As little as possible should stand between the storyteller and the audience.

ROM; and a programming universe that includes picture-making tools as well as all the usual database, word-processing, and file-handling engines. Only with all this present is it safe or worthwhile to add the narrative-modeling system (that as yet hardly exists).

Who is the author?

There remains a further problem: Exactly who is going to use such a system? One thing is certain—it must not be a programmer. Just as a ghost-written book only pleases the gullible and a TV science documentary devised by a professional broadcaster is only for entertainment, so an IEP not created by the real owner of the knowledge can be fatally flawed: It has gone through the filter of someone else’s perceptions of the problem. The distinction here is like that between the literary editor, who can enable the author to improve his or her own output, and the film producer, who uses the author’s work as input only weakly coupled to the output.

As little as possible should stand between the storyteller and the audience. This puts the heaviest burden of all on the creation workstation. It must support the activities of a non-programmer and still make the full resources of the system available both for trial and construction. It is most unlikely, therefore, that such an engine can appropriately *deliver* IEPs.

Putting the user first

IEPs focus primarily on the intended user, not the creator. Users will outnumber creators by a wide margin. The idea that today we are all authors is no truer now than when the pencil

was invented. The new book, with its expansion into interactive control of moving pictures and sound, can be the enabling technology for the millenium. I argue that this technology is a special-purpose engine rather than an extension of the multipurpose PC.

The delivery engine

The demands of an IEP delivery engine differ markedly from those of the workstation used for its construction. While the construction process is initially open ended and experimental, the use of the finished product is deterministic and finite. That does not mean that to the end user the system appears closed and finite. On the contrary, good IEPs must present as open and exploratory a face as possible. Nevertheless, this appearance of openness is an illusion. All the pathways are latent and all the interconnections exist—everything is there, just waiting to be uncovered. This essential determinism simplifies the demands on the operating system. What now becomes essential is the cheap and efficient manipulation of runtime images and sounds.

With one or two honorable exceptions,³ the systems so far devised to present images under interactive control have only been used as “or” devices. The storage system has been used to deliver this image or that image or the other image because, historically, the image store itself supports the video display, thereby removing the possibility of having more than one image at once. This profoundly inefficient way of going about things carries over from the days when computers found video not just very difficult but impossible to support.

The need for comparison

At the heart of all acts that combine to permit the revelation of a new idea lies the act of comparison. Surely we never wish to see a picture in isolation—we always require the juxtaposition of images.

We have the power to commute time and space to some degree. The development of a series of images in time, as in a movie, has a cumulative effect, and the change from each image to the next is the nexus of meaning. Unfortunately, this only works for the limited changes in response to which our visual system has evolved. If the two temporally adjacent images are too dissimilar, we perceive them as unrelated and often, in consequence, reject them as incomprehensible.

For educational efficiency, sequential presentation is rarely adequate. The images to be compared must appear simultaneously. Clearly, a prime constituent of an IEP delivery system is some kind of frame store that can take more than one image from the image store and present them in educationally appropriate ways.

Storage and manipulation

It seems logical now to build appropriate functionality into the image storage devices themselves, rather than to treat such

devices as peripherals to a computing engine. In the very earliest days of interactive video this point was already well understood, and devices characterized as Level 2 in the Nebraska taxonomy⁴ formed a key sector in the early markets. Unfortunately, development moved away from this field to the so-called Level 3 systems, in which heroic efforts were made to incorporate analog video into the computer milieu.

It is time to return to the Level 2 concept. Rather than designing additional boards to slot into the various hardware platforms so that videodisc and other image media can be incorporated into the computer environment, we should design ex-

***Surely we never wish to see a picture
in isolation—we always require the
juxtaposition of images.***

tensions to the videodisc and other players appropriately to enhance their functionality.

Most of the major videodisc drive manufacturers now support rudimentary character generation. An international standard covers the mixed-mode laser disc, which makes analog video and digital data in CD format simultaneously available. Commercial videodisc players already on the market contain field and frame stores. It should be possible to merge the undoubted strengths of digital audio and runtime graphics with the proven versatility of the analog videodisc.

Superior video systems

A number of strategic improvements are possible on the video side. The most important of these involves replacing the existing composite video format (NTSC or PAL) by the Multiplexed Analog Components (MAC) system.⁵ This encoding scheme has two great advantages. First, it removes all the deleterious effects of the color subcarrier and paves the way for extensions to HDTV via versions of the Muse⁶ (Multiple Sub-Nyquist Encoding) system. Second, the chip set designed to implement the system recognizes up to 256 distinct uses for the 52 microseconds of active line time. In this way, therefore, even successive lines in a video raster can be used to represent different classes or types of data. The most important consequence is that entire video fields can be used to store digital data in addition to the separate digital data channel carrying CD-DA or CD-ROM.

Integrating video

The argument presented above, which develops the essential requirement that any such system must function in a comparative mode for images, implies that the delivery engine must contain a frame store supplied both from the digital channel of

the disc and, via A/D conversion, from the analog channel. The panoply of techniques from the image processing field can now be delivered in silicon. This makes operations such as resizing and compositing available to the user as runtime operations for display (under program control) of arbitrary images from the disc in pairs or other composites. It seems probable that the speed of access conferred by constant angular velocity (CAV) disc strategies is preferable over the increase in capacity obtainable by constant angular acceleration (CAA) techniques, even though the discs are likely to measure eight inches or less in diameter rather than the full 12 inches available to the laser disc format.

Integrating sound

To date, the area that has received the least attention concerns the ways in which the various sources of sound can be integrated into such a system. The film industry has long recognized several classes of sound, for example, lip synch, voice-over, spot effects, wild tracks, and background fills. It might be wise to incorporate into this repertoire the possibility of recording, comparing, and reproducing real-time input from the user. It is in this area that the significant advances implemented in desktop computers could be carried over to the proposed new delivery engines. The ability to provide many different channels of speech-quality sound over a programmable succession of still images will become one of the most common presentation styles in IEPs. The fact that the laser disc format allows the storage of CD-DA data in addition to analog video favors that format significantly.

A uniform device

Perhaps the best way to regard a delivery engine of the type indicated above is to think of it in classical device terms. That is to say, it has a set of internal functional possibilities made available to the user through an instruction set. This strategy implies that many manufacturers, provided they delivered identical functionality, could choose to deliver that functionality in various ways transparent to the end user. However, it is essential to have an agreed standard platform on which all authors could publish.

Delivering an IEP on such a system entails delivering a data structure and a program whose operation codes are the instructions that the device driver recognizes. In conventional software terms this is rather like executing a runtime kernel. But, whereas conventional runtime systems call on a subset of the functionality of the same engine that created them, in this case the runtime code targets a specific, independent, free-standing device.

A single box

A key strategy, already presaged by some manufacturers, concentrates development on a compact single-box solution that incorporates the display. This strategy arises because the

new device must be no less convenient than today's book. There remains the question of the most suitable input device for user interaction. The point-and-click metaphor, especially when it can be generalized from constructed entities such as icons and hypertext to arbitrary images,⁷ seems most likely to be adequate for the perceptual path of the feedback route. Nonetheless, questions remain as to the relative merits of trackballs, scratchpads, and other such space-constrained devices.

Learning from CD-I

Many of these problems are being addressed by those engaged in the development of compact disk-interactive (CD-I), and we can learn much from their experience. It is safe to say that we know far too little about the environment necessary to support the creation by nonprogrammers of an effective interactive narrative. It seems likely that a fully object-oriented universe will offer the level of flexibility and real-time intervention essential in the creation of the multilayered webs needed to encapsulate the ideas in an interactive book that responds differently to different "readers."

It is important to realize, however, that the greatest stabilizing

An international standard covers the mixed-mode laser disc, which makes analog video and digital data in CD format simultaneously available.

feature in the design of authoring environments is the precise definition of the functionality of the target delivery device. Once these properties are established from pedagogical and artistic criteria, the technological and programming issues will fall into place.

Perhaps the most important matter conditioning the acceptance of CD-I is the significance of full-screen, full motion (FSFM) video. Because it is so difficult to achieve FSFM images if the data rate of the storage device meets the standard for CD-DA (as it must for manufacturers to exploit existing economies of scale), the current developers make a virtue of necessity. Unfortunately, it might well be that consumers must see interactivity as a smooth upgrade path from existing videotape presentations. If so, the FSFM is essential. That laser disc can deliver this and still have the ability to use digital data is a major factor in favor of that format.

A parallel

We can make an interesting comparison with the last major event in personal computing, namely desktop publishing. In my view, the overriding factor in the success of DTP has been the

IMPORTANT PROCEEDINGS

FROM



**IEEE Computer
Society Press**

COMPUTER GRAPHICS AND DESIGN

1990 IEEE INTERNATIONAL CONFERENCE ON COMPUTER-AIDED DESIGN (ICCAD-90)

598 pages. ISBN 0-8186-2055-2.

Catalog No. 2055 List \$80.00 Member \$40.00*

1990 IEEE CONFERENCE ON COMPUTER DESIGN: VLSI IN COMPUTERS AND PROCESSORS (ICCD '90)

498 pages. ISBN 0-8186-2079-X.

Catalog No. 2055 List \$98.00 Member \$49.00*

1990 IEEE WORKSHOP ON VISUAL LANGUAGES

258 pages. ISBN 0-8186-2090-0.

Catalog No. 2090 List \$75.00 Member \$35.00*

27TH ACM/IEEE DESIGN AUTOMATION CONFERENCE

766 pages. ISBN 0-8186-9650-X.

Catalog No. 2050 List \$130.00 Member \$65.00*

IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR '89)

714 pages. ISBN 0-8186-1952-X.

Catalog No. 1952 List \$120.00 Member \$60.00*

**All prices are 20% higher outside the U.S. and Canada*

SEND YOUR ORDERS TO:

IEEE COMPUTER SOCIETY PRESS
10662 LOS VAQUEROS CIRCLE
P.O. BOX 3014
LOS ALAMITOS, CA 90720-1264

Or Call Toll-Free
1-800-CS-BOOKS

Or in California
714-821-8380

Or FAX **714-821-4010**

(ADD \$5.00 PER BOOK TO COVER HANDLING CHARGES.)

existence of the Apple Laserwriter—the delivery engine for desktop publishing. When launched, the Laserwriter was intrinsically more highly resourced in computational terms than the majority of the computers that accessed it. Apple designers had carefully tailored the printer to undertake a specific set of operations. They had made it as fit for the purpose as possible by augmenting the mechanical printing engine with a powerful processor, a large quantity of RAM, and a ROM containing a highly specialized device driver, written in a purpose-built page description language. The success of the Laserwriter in producing sheets of paper that contain the marks carefully assembled by a variety of text manipulation programs has been remarkable. This success results from its optimization for the purpose. Exactly the same state of affairs must hold for a revolution in the way in which we impart knowledge and tell stories using the interactive electronic medium.

Recommendation

The time is ripe for an agreement as to the appropriate functionality that the new “book” must possess and for these properties to be embodied in a cheap, robust device that can take its place in the home, the workplace, and the school. It is time that some of those people with the requisite skills divert their concentration away from increasing the scope of computers, which only a minority will ever own, and concentrate their talents on increasing the functionality of end-user devices that can revolutionize the way future generations understand the world in which they live. □

References

1. M.G. Carasso, J.B.H. Peek, and J.P. Sinjou, “The Compact Disc Digital Audio System,” *Philips Technical Review*, Vol. 40, No. 6, 1982, pp. 151-164.
2. S. Brand, *The Media Lab—Inventing the Future at MIT*, Penguin Books, London, 1989.
3. W.E. Mackay and G. Davenport, “Virtual Video Editing in Interactive Multimedia Applications,” *Comm. ACM*, Vol. 32, No. 7, 1989, pp. 802-810.
4. R. Daynes, “The Videodisc Interfacing Primer,” *Byte*, Vol. 7, No. 6, 1982, pp. 48-58.
5. R. Morcom and G. Drury, “MAC—The Video Coding System for DBS,” *J. British Kinematic Screen and Television Society (BKSTS)*, Vol. 65, No. 11, 1983, pp. 602-611.
6. T. Toyama et al., “Optical Videodisc for High-Definition Television by the Muse,” *SMPTE J.*, Vol. 95, No. 1, 1986, pp. 25-29.
7. D.R. Clark and N. Sandford, “Semantic Descriptors and Maps of Meaning for Videodisc Images,” *Programmed Learning and Educational Technology (PLET)*, Vol. 23, No. 1, 1986, pp. 84-90.



David R. Clark has been Director of the University of London Audio-Visual Centre since 1985. He has worked in remote sensing since 1978 and in computer graphics since 1971. He took his PhD at Manchester University in explosion theory in 1966 and taught at the University of Toronto until 1971, when he started a new career as a film maker, working with Samuel Beckett and the Royal Shakespeare Company and making many educational programs.

Clark can be reached at North Wing Studios, Senate House, Malet St., London WC1 7JZ, UK; on Applelink UK0045 or via Janet as drc@uk.ac.bbk.cs.

Jim Blinn's Corner



It turns out that the obvious way to transform NDC to pixel space is wrong. To show you why, I'll have to go through several intermediate stages before I get to the correct way.

A Trip Down the Graphics Pipeline: Pixel Coordinates

James F. Blinn, Caltech

There are some problems in graphics that I've spent a lot of time trying to figure out even though they seem real simple. I don't know if this is because I'm dumb or because these problems are really profound. I've never seen any of them described anywhere in the literature; maybe the people who write books just haven't come across them.

One set of such problems comes from the use of transformations in the graphics pipeline. No, I'm not going to beleague you with yet more ravings about how wonderful homogeneous coordinates are. You already know that. (And if you don't, you might want to refer to the Big White Book.¹) What I am going to do is discuss a few of the odd little quirks of the transformation process that I've discovered over the years and how I learned to deal with them. Some of this may look like I'm making a big deal out of nothing, but it's important to get it right.

More specifically, over the next few issues I am going to talk about

1. a detailed examination of what it means to be a pixel,
2. a nifty way to merge the old concept of the window-to-viewport transformation with clipping space that makes it easy to clip viewports, and
3. what's really going on with the homogeneous perspective transform.

These are part of a series begun with my column in the January 1991 issue ("A Trip down the Graphics Pipeline: Line

Clipping"). For variety, I probably will intersperse these discussions with columns on other topics.

Overview of the pipeline

Let's start with a general overview of the classic rendering pipeline. Basically the pipeline transforms points (line endpoints, polygon vertices, patch control points, and so forth) from some object's definition space to pixel raster space and, along the way, clips the shapes to the boundaries of some given region of space. The pipeline is first *initialized* by setting up various global parameters (primarily transformations), and then it is *used* by passing points through it. There are two ways to look at the setup process: from the user's *external* point of view and from the pipeline implementor's *internal* point of view. The user specifies the parameters that describe what he/she/it wants. The pipeline implementor figures out how they are stored and used by the pass-through process.

Coordinate Spaces

First let's look at the user's view. There are various coordinate systems that a point experiences as it journeys to the screen. In the interests of solidarity I'll try to relate my notation to that used by Pixar's Renderman interface, which I will refer to as the politically correct Renderperson. In the following list the name used by Renderperson appears in parentheses. I'll list the coordinate spaces in the order in which an object sees them.

1. Definition (object)—The space in which the original object is defined.
2. Universe (world)—A consistent universe into which all modeled objects are placed.
3. Eye (camera)—Eye at origin looking down the z axis.
4. Perspective—Distorted space after the perspective transform has been performed. This one is optional. If there is an orthographic projection, this will be the same as eye space.
5. Clip—A special space chosen to make the clipper's arithmetic simple. It maps the desired visible chunk of perspective space to a fixed region.
6. Normalized device coordinates (which I will refer to as NDC) (screen)—A standardized screen coordinate system that doesn't include explicit pixel dimensions.
7. Pixel (raster)— X, Y coordinates (usually integers) in the hardware pixel space. In modern usage, with multiple screen windows, this is the screen real estate we have been allocated by the windowing system.

Here are some notes about the transformations between these coordinate systems (and advertisements for back issues of this magazine):

Definition to Universe is generally a combination of rotation, translation, and possibly nonuniform scales used to squash primitives to the desired shape and place them in the universe. It is sometimes something more bizarre, like the bend-me, twist-me, hurt-me transforms used in general deformation systems. The ability of a 4×4 matrix to do perspective is rarely used in modeling except in rational polynomial curves and surfaces or local light source shadows. I discussed the latter in January 1988 (CG&A, "Me and My (Fake) Shadow").

Universe to Eye is just a translation and pure rotation to place and orient the simulated camera. I wrote about this in the July 1988 column ("Where Am I, What Am I Looking at").

Eye to Perspective is a homogeneous perspective transform, the subject of an upcoming column.

Perspective to Clip to NDC is the subject of another upcoming column that fills in the blanks from my January 1991 effort ("Line Clipping").

NDC to Pixel is the subject of most of this column and its sequel.

Internal operation

Now let's look at internals. Many of the above coordinate systems exist only in our minds. In actual implementation all the transforms from definition to clip space are merged into one transform which we'll call T_1 . This is usually a full 4×4 homogeneous matrix. The transforms from clip to pixel space are merged into T_2 . This usually requires only a scale and offset in X and Y .

Once the transforms have been set up, processing an object through the pipe requires four steps:

1. Transform to clipping space via T_1
2. Clip

3. Divide by w
4. Transform to pixel space via T_2

Exemplary transformations

The only transformation I will need for the next few columns is a simple scale and offset in X and Y from some input coordinate system to an output coordinate system.

$$s_x X_{in} + d_x = X_{out}$$

$$s_y Y_{in} + d_y = Y_{out}$$

It is easiest to specify these as exemplary transformations. No, this doesn't mean that they are more admirable than others. It just means that they are specified in terms of two *example* input coordinate values and their associated desired output values. I'll denote this for, say, X as

$$X_{in1} \mapsto X_{out1}$$

$$X_{in2} \mapsto X_{out2}$$

Using the magic of linear transformations, we can derive the scale and offset as

$$s_x = \frac{X_{out2} - X_{out1}}{X_{in2} - X_{in1}}; \quad d_x = X_{out1} - s_x X_{in1}$$

These values are what is actually stored for use during the pass-through stage. For the transforms derived below, I'll just use these equations and show you the results.

NDC to pixel transformation

Let's start with what might at first seem the simplest transformation: NDC to pixel space. The transform is

$$s_x X_{NDC} + d_x = X_{pixel}$$

$$s_y Y_{NDC} + d_y = Y_{pixel}$$

A user/programmer does all screen design in NDC. There are three nasty realities of the hardware that NDC hides from us:

1. The actual number of pixels in x and y .
2. Nonsquare pixels. Distances in NDC space are geometrically uniform.
3. Up versus down for the Y coordinate. It will invert Y if necessary so that Y in NDC points up.

The transformation to pixel space is implicit in the system and is set up during device initialization. It turns out that the obvious way to transform NDC to pixel space is wrong. To show you why, I'll have to go through several intermediate stages before I get to the correct way. But without all the false starts the correct way will seem kind of unmotivated. So don't take any of these stepping stone transformations too much to heart until I tell you we've arrived at the ultimate one.

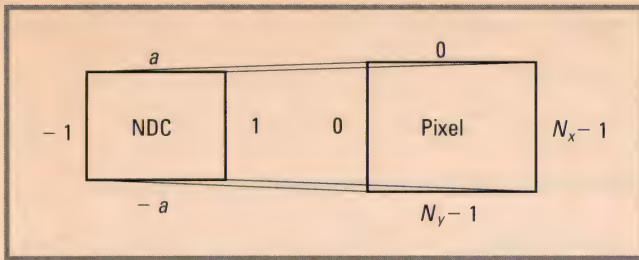


Figure 1. The naive NDC to pixel model.

First let's specify precisely what I mean by the different coordinate systems.

Pixel space

Define the pixel coordinate system according to the following typical scheme. The pixel array is N_x pixels across, numbered from 0 on the left to $N_x - 1$ on the right. The array is N_y pixels high, numbered from 0 on the top to $N_y - 1$ on the bottom. Now, in addition to deriving general formulas, I'm going to use explicit values in examples below. I'll use 512 pixels in X and 480 pixels in Y , not just because they are typical but also because they are familiar and you'll be able to see what's going on with them.

Sometimes the pixel array doesn't start at (0, 0) but at some offset. I won't specifically deal with this here; it only requires adding the constant (X , Y) offset to the pixel coordinates we calculate. Also, sometimes $Y = 0$ at the bottom of the screen, instead of at the top. I'll let you figure out the slightly different transform necessary in that case.

Now in addition to pixel counts the display hardware will imply a physical aspect ratio for this rectangle of pixels that might not necessarily equal N_y/N_x . Thus, pixels might not be square. For our numerical example, I'll use a standard TV screen which has an aspect ratio of $3:4 = 0.75$. This number represents the height divided by the width of the region that the display electronics will scan out. For some reason Renderperson describes the aspect ratio as width over height.

At some level it might not make any difference. I think the aspect-ratio-sensing neurons of the general public are getting burned out because so many computer displays are not really adjusted to the proper hardware aspect ratio and are distorted horizontally. In addition, people see so many images on TV that are distorted via digital video effects that pretty soon they don't even notice it anymore. This is similar to how my flicker sensors have been burned out from my early years of looking at displays that refreshed only about 10 times per second.

NDC space

Normalized device coordinates run from -1 to $+1$ in X and $-a$ to $+a$ in Y , where a is the above mentioned physical aspect ratio. NDC thus has a little bit of device dependency (see the left side of Figure 1).

The Renderperson approach is perhaps a bit better. It guarantees that the region -1 to $+1$ is visible in both X and Y ; the longer dimension then extends to values beyond 1. They would therefore say that TV screen space extends from -1.333 to $+1.333$ in X and from -1 to $+1$ in Y . I'll stick to my way for now so I don't get confused.

The naive transform

In Figure 1 we see our first attempt at the required transformation. Let's derive the exemplary transformation from this figure. In X

$$-1 \mapsto 0$$

$$1 \mapsto N_x - 1$$

This leads to

$$s_x = \frac{N_x - 1}{2}; \quad d_x = \frac{N_x - 1}{2}$$

For our sample pixel range this leads to

$$s_x = 255.5; \quad d_x = 255.5$$

Now we do the same thing for Y . Here we incorporate the physical aspect ratio of the screen, a , as the vertical bounds for NDC. This implicitly corrects for nonsquare pixels.

$$-a \mapsto N_y - 1$$

$$a \mapsto 0$$

so

$$s_y = \frac{N_y - 1}{-2a}; \quad d_y = \frac{N_y - 1}{2}$$

For our sample pixel range this leads to

$$s_y = -319.333; \quad d_y = 239.5$$

Every computer graphics book I've seen (at least every one that describes this explicitly enough to tell) uses this transformation.

A crinkle

There's a problem lurking here, though. The scale factor is icky. In the X direction we are taking two units of NDC space and dividing it up into N_x pixels. We sort of expect the scale factor to be $N_x/2$. What did we do wrong? We forgot that pixels have size. Imagine a row of pixels with a Gaussian spot centered at each one and you realize that the resultant intensity swath sticks out about $1/2$ pixel width to the left and right.

This little offset doesn't seem like much, but it can really screw things up for more advanced rendering situations. I did this wrong in my original teapot pictures.² There, due to memory limitations, I had to make each 512×512 -pixel image from four 256×256 -pixel quadrants using a variant of the naive pixel mapping. For example, for the left two quadrants, the X mapping was $-1 \mapsto 0$ and $0 \mapsto 255$, and for the right

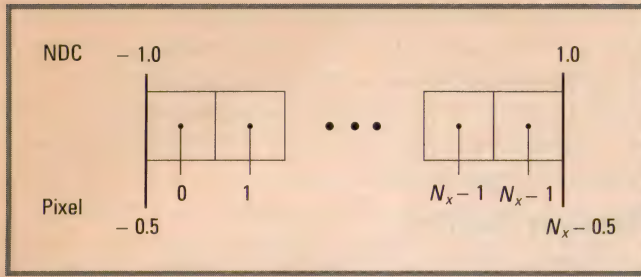


Figure 2. Microscopic view of pixels at the edge of NDC.

two quadrants, it was $0 \mapsto 0$ and $1 \mapsto 255$. When the four pictures were composited, you could see a seam. (This was only really noticeable when the digital image was zoomed up—another example of the principle that if you shrink a crummy picture down, it looks good.)

So what's the right way? Looking at the pixels through a microscope, what we *really* want is shown in Figure 2. For the X coordinates, we use

$$-1 \mapsto -0.5$$

$$1 \mapsto N_x - 0.5$$

so

$$s_x = \frac{N_x}{2}; \quad d_x = \frac{N_x - 1}{2}$$

With explicit coordinates we get

$$s_x = 256; \quad d_x = 255.5$$

For the Y coordinate, we use

$$-a \mapsto N_y - 0.5$$

$$a \mapsto -0.5$$

so

$$s_y = \frac{-N_y}{2a}; \quad d_y = \frac{N_y - 1}{2}$$

and

$$s_y = -320; \quad d_y = 239.5$$

Integerizing

When we go from floating-point pixel coordinates to integer pixel coordinates, we need to round to the nearest pixel. This is the same as adding 0.5 and truncating. We might as well build this into the displacement part of our transformation. This implies an X map of

$$-1 \mapsto 0$$

$$1 \mapsto N_x$$

and a Y map of

$$-a \mapsto N_y$$

$$a \mapsto 0$$

This changes just the displacements to

$$d_x = \frac{N_x}{2} = 256$$

$$d_y = \frac{N_y}{2} = 240$$

This works as long as all the numbers are positive, which they are here.

The dangling edge

Now there's another problem. An NDC X coordinate value of 1.0 will map, after truncation, into N_x , which is one pixel beyond the right edge of the screen. There are two ways we can react to this situation. First, we can disallow the value of +1.0 as a visible coordinate, that is, fix it so the clipper keeps things only in the region $-1 \leq X < +1$. But this leads to some unexpected and undesirable results. For example, if you try to frame the screen by lines tracing what you think are the screen boundaries (-1 and $+1$), you won't get to see the line on the right.

Second, we can arrange to keep this one extra NDC value in a way that won't mess us up. Set up the clipper to keep both ends: $-1 \leq X \leq +1$. We can then test for the pixel coordinate of N_x and explicitly set it back to $N_x - 1$, but this is not what I do. I do what all good computer graphicsists do, I cheat. I contract the NDC space by a microscopic amount to fit the extra one value into the allowable pixel range.

In X ,

$$-1 \mapsto 0$$

$$+1 \mapsto N_x - \epsilon$$

This will make $+1$ truncate to $N_x - 1$. The scale and displacement are

$$s_x = \frac{N_x - \epsilon}{2}; \quad d_x = \frac{N_x - \epsilon}{2}$$

You can use numerical analysis to figure out the optimal value for ϵ but I just use 0.001, which is pretty close to the minimum value whose effect still fits in a single precision number.

$$s_x = 255.9995; \quad d_y = 255.9995$$

In Y ,

$$-a \mapsto N_y - \epsilon$$

$$+a \mapsto 0$$

so

$$s_y = \frac{N_y - \epsilon}{-2a}; \quad d_y = \frac{N_y - \epsilon}{2}$$

and

$$s_y = -319.9995; \quad d_y = 239.9995$$

This, then, is the final approved NDC to pixel transformation. Write this down.

What is pixel space really?

So what have we done here? Considering just the X direction, we've divided the conceptually continuous -1 to $+1$ range of NDC up into N_x discrete bins, each of which maps to a pixel. There are actually two ways to think about this. We can place the integer pixel coordinate at the center of the pixel and then round each coordinate to the nearest integer. (This was one of our intermediate techniques.) Alternatively, we can place the integer pixel coordinate in the cracks between pixels and truncate to the next lowest integer. This is what we finally wound up with. Note that the pixel areas are geometrically the same for both techniques. The left-most pixel covers $2/N_x$ amount of distance in NDC, abutting the left edge of the screen.

Simple line drawing algorithms and many early polygon scan conversion algorithms accept line endpoints and polygon vertices as integer pixel coordinates. That's great for quick-look line drawings, plotters, or other output devices that take integer pixel coordinates. But there's more to life than line drawings.

Mulling over some more advanced rendering applications leads to two other requirements of the NDC to pixel mapping: subpixel positioning of endpoints/vertices and subsampling of the image for antialiasing purposes. These two enhancements to the transformation interact in strange and mysterious ways and will be the topic of next issue's column. □

References

1. J.D. Foley et al., *Computer Graphics: Principles and Practice*, Addison-Wesley, Reading, Mass., 1990.
2. J. Blinn, "Texture and Reflection in Computer-Generated Images," *CACM*, Vol. 19, No. 10, Oct. 1976, pp. 542-47.

Post mortem on the last issue

I was pretty happy about the appearance of the diagrams from last issue's column about printing quality ("WYSBOAVRTWYG"). I was especially happy that *all* my very thin lines survived the printing process. It did look like the 0.08 (2.8 pixel) and 0.04 (1.4 pixel) point lines were the same width, though. The halftone lines didn't look bad either, despite my worries about line position interacting with the halftone pattern. You *can* see the effect I was talking about in Figure 2 by looking at Figure 3 through a loupe (that's a small magnifying glass for those of you who don't speak French).

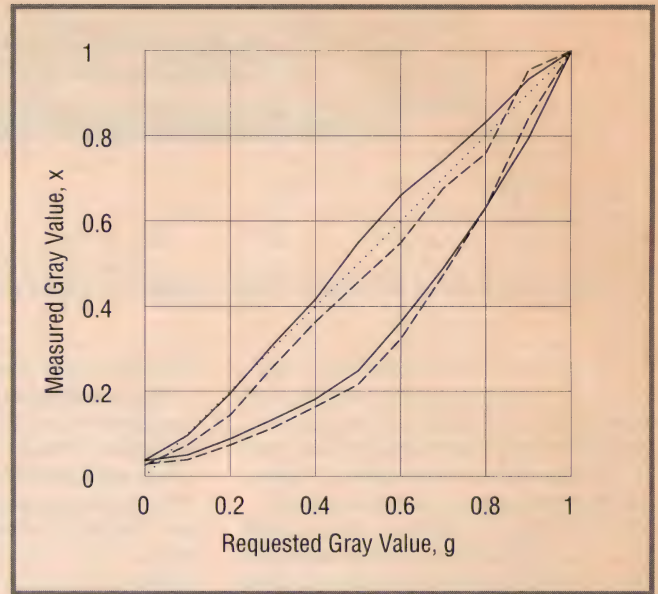


Figure 3. Measurement of compensation of CG&A Postscript-to-print distortion.

The compensation table example in Figure 14 looked a lot less murky than the uncompensated version in Figure 8. Because of the magazine's layout, you can easily compare the figures side by side by folding the page over. Finally, as promised, Figure 3 is a graph of the measured gray values of the compensated wedge (which should ideally be an identity function) compared to the uncompensated wedge. The dotted line represents the ideal identity function. On top of this I have plotted measurements of the original and compensated wedges as measured by two different densitometers (one meter is the solid lines and the other is the dashed lines). Using two meters gives you an idea of how accurate a densitometer reading is likely to be. The two sagging lines are from the uncorrected wedge. The two straightish lines are from the corrected wedge, not too bad a result, considering that the original data was printed almost six months before I used it to make the compensation function.

Editor's Note: Jim Blinn recently received a MacArthur Fellowship for his work in education and computer graphics. Awarding Blinn a grant totaling \$265,000, the John D. and Catherine T. MacArthur Foundation stated that, "in designing computer graphics, he works simultaneously as an artist, educator, mathematician, and scientist." Blinn is among the 31 scientists, scholars, and artists who have received this prestigious award this year. This news reached our offices right before press time. More complete coverage will appear in our next issue.

Selective Update

Computer Graphics and Applications seeks new logo

IEEE Computer Graphics and Applications is holding a contest to find a new computer-generated logo (the graphic presentation of the magazine's name as it appears on the cover). *CG&A*'s first ever computer graphics contest challenges computer graphics professionals and students to show off their skills and the capabilities of their art, design, drawing, or publishing software.

The redesign is part of *CG&A*'s continuing efforts toward updating its appearance and streamlining its

production with electronic publishing. *CG&A* is currently produced on PCs that run Ventura Publisher, Corel Draw, and other software.

The old logo was drawn by hand by artist Jay Simpson of Design and Direction. Simpson's design was then scanned into an Apple Macintosh IIfx by independent artist Alex Torres, who refined the logo with Aldus Freehand software. The polished design was then converted to an encapsulated Postscript (EPS) file to be imported into and manipulated in

Ventura Publisher.

But with *CG&A*'s recent anniversary, the logo turns a respectable (if well-worn) 10 years old. Software and hardware have obviously improved over the past decade, and so has *CG&A*'s willingness and ability to take advantage of the new technology. Thus, *CG&A* seeks a computer-generated logo that more accurately reflects the trends in graphics and publishing in the nineties.

For an explanation of how to enter and an entry form, see pages 89 and 90.

Siggraph 91 employs multimedia in digital invitation

ACM Siggraph 91 took advantage of developing imaging technology to send out multimedia digital invitations for this year's conference, scheduled for July 28 through August 2 in Las Vegas, Nevada. Recipients of the disks can browse the activities planned for the 18th conference on computer graphics, ranging from the exhibit floor through course sessions to art objects.

Dick Mueller of Dynamic Graphics, Minneapolis, served as executive producer and art director of the digital invitation. He said Siggraph 91 put to work every bit of technology available for imaging to produce the invitations. They used Hell scanners (for the logo), Ras-

ter Ops monitors and frame grabbers, and 35mm scanners. Authorware donated software. The disks run on color Macintoshes or DOS-Window machines that support 256 colors.

Mueller asserted that the technology used is the forerunner of the next generation of advertising. He said it costs less to produce the interactive disks than to create a full-color brochure, and programming takes less than 30 days.

For a copy of the digital invitation or for more information about the conference, contact Siggraph 91 Conference Management, 401 N. Michigan Ave., Chicago, IL 60611, phone (312) 644-6610, fax (312) 321-6876.

European design automation conference planned

A European conference and research exposition patterned after the Design Automation Conference is planned for September 7-11, 1992, at the Hamburg Messe and Congress Center in Hamburg, Germany. Euro-DAC 1992 is chaired by Joseph Borel of SGS Thomson, France, and owes much to the efforts of Egon Horbst, Siemens AG and the European/Middle East representative on the DAC executive committee.

Euro-DAC consists of a technical program, tutorials, and vendor exhibits. Program chair Gerry Musgrave of

Brunel University, UK, said the technical program will focus on design automation technologies for design, test, engineering, and manufacturing with an emphasis on the user's viewpoint.

For detailed information, contact Euro-DAC, Siemens-Nixdorf Informationssysteme, Anton Sauer, 5 NI AP 27, Rosenheimerstr. 116 b, D-8000 Munich 80, phone (49) 89-4111-2680, fax (49) 89-4111-2685, or Euro-DAC, MP Associates, 7490 Clubhouse Rd., #102, Boulder, CO 80301, phone (303) 530-4562, fax (303) 530-4334.

INEL extends invitation to gallery exhibitors

The Idaho National Engineering Laboratory (INEL) has scheduled its fifth annual computing symposium for September 10-12, 1991 at the Shilo Inn in Idaho Falls, Idaho. The gallery exhibit accompanying the symposium invites participants to submit computer-generated graphics from scientific visualization, computer-generated animation, and multimedia. The gallery will focus on the use of computers for visualizing data and as a medium for creative interpretation.

Submissions must be received by August 27, 1991. Work produced between August 1990 and August 1991 is eligible. Formats should be 16 x 20 glossy color print (mailed flat, not bent or folded), 35mm color slides, or SVHS or VHS video tape. Submissions will not be returned unless the entrant specifically requests it.

Credits should include the creators, laboratory or company name, hardware and software, and intended end use.

Mail submissions to Nelson Soucek, EG&G Idaho, PO Box 1625, Idaho Falls, ID 83415-2608, phone (208) 526-9321, fax (208) 526-9936.

The symposium reportedly provides a forum for participants to discuss the application of computers, supporting computer technologies and techniques, and future directions.

NCGA 91 does windows in Chicago

Tom Culviner, Contributing Editor

The National Computer Graphics Association's 12th annual conference, held April 22 to 25 at Chicago's McCormick Place, attracted approximately 21,000 people, who came to learn about the products and services of 180 exhibitors and attend 86 conference sessions. The theme of the conference, "We're Doing Windows," was reflected in sessions and demonstrations showing uses of windows-based graphical interfaces.

Ultimate user

Richard Leet, vice chairman of Amoco Corp., gave the keynote address, "Harnessing the Information Explosion: A Businessman's Perspective." Leet prefaced his talk with the observation that he was not a computer graphics expert; rather, he represented the "ultimate user," a person whose sole connection with computer graphics is to see and use the final results.

Leet reviewed some of the milestones in Amoco's growth as an oil refining company and outlined how Amoco uses computers and graphics in computational chemistry, simulation, and the processing of satellite data. According to Leet, the increased information brought on by computerization presents a danger. An organization might find itself gathering and retaining large amounts of information simply because technology lets people do this.

Warning the audience that "information is not knowledge," Leet said that as hardware improves, computers aim in increasing amounts of information at us. But we cannot plug in additional memory and processing capabilities to handle the information. Computer graphics is important because it presents all this information in a form that lets us work productively with it. According to Leet, what users want from computer graphics in the future are faster machines—no wait time after asking "what if?"—and systems that can display larger data files in one picture.

Orr goes first

NCGA is not complete without hearing what Alan Paller, of AUI Data Graphics/CA, and Joel Orr, of Orr Associates, have to say about the future of computer graphics. At past NCGA sessions, Paller spoke first—quickly and confidently—about specific products and developments, then Orr used a

larger view to refute patiently and methodically much of what Paller had said. This year the formula changed, and Orr went first.

Without Paller's setups, Orr could not apply his broader view to specifics the audience had just heard from Paller. Nevertheless, this was one of NCGA's most interesting sessions.

Orr observed that in our time technology drives everything, just as in the last century Manifest Destiny drove the U.S. As a result, we make devices to perform functions whose implications we can't understand, and by the time we understand the implications of the first phase, the devices are doing what they do twice as fast.

Said Leet, what users want from computer graphics in the future are faster machines—no wait time after asking "what if?"—and systems that can display larger data files in one picture.

A consequence of such rapid change is that the one-year lag in publishing a textbook about a technological subject is no longer acceptable. By the time the book is published, it's obsolete. Another consequence is that software has many more capabilities than most users are aware of. A major challenge is to put the capabilities on "the surface" without cluttering the surface. Based on information from cognitive psychology—that for us to feel intimate with something, it must fill a large part of our visual field—Orr said that larger computer screens would result in better user interfaces.

Orr predicted that the 1990s will see manufacturers coming to terms about standards. As machines get smaller and cheaper, manufacturers have to sell more of them. Manufacturers also must satisfy purchasers who want to be sure that what they buy will work with what they already have or might get later.

Orr also predicted that in this decade the graphics board industry will virtually disappear. Very high resolution processing will be available on the

motherboard, so separate graphics boards will be needed only for a few special applications.

Paller followed Orr with terse predictions and observations, including the following: Mainframe and minicomputer graphics is dead, and the main growth will be in graphics on Unix terminals; Harvard Graphics has won in its PC application niche because it is convenient to use; Microsoft is winning the windows war, even against the Macintosh; and every Macintosh software vendor has moved or is moving its software to the DOS-windows environment.

Paller predicted that soon all data will be stored on video disks—once fast image-compression technology matures. The short-term results will be better presentations. In the long term, we can expect total document recall and total meeting recall, so instead of reading minutes for a past meeting, you will scan and view the actual meeting. A related consequence will be that all visuals for an upcoming meeting will be kept on line and up to date. You will be able to review the visuals, decide whether you want to attend the meeting, and skip it if you don't.

Visualization and more

At his session "Where Visual Computing is Going in the 1990s," Carl Machover identified the financial industry as an area where visualization could be applied to help analysts interpret large amounts of data. Machover said that the payoff for an implementation could come in one successful trade, instead of in several years, as is normal in other industries. According to Machover, companies already trying financial applications are reluctant to talk about what they are doing, because their implementations give them a competitive advantage.

The Graphics Performance Characterization Committee released beta results from application test files for the Picture-Level Benchmark. The first issue of *The GPC Quarterly* (published by NCGA) reports results for 25 different workstations running five different applications files. For the test, an application file must be converted to the benchmark's Binary Interface Format. Then the file can be run on any hardware platform that has a Picture-Level Benchmark port. Thus, application results are standardized for any hardware platform on which the file is run. The re-

port categorizes according to price and gives results for workstations from Digital Equipment, Evans and Sutherland, Hewlett-Packard, IBM, and Sun Microsystems.

At its conference reception, NCGA presented its annual achievement awards. Michael Boich, president and CEO of Radius, received the executive of the year award. Karl Gutttag, graphics strategy manager at Texas Instruments, and Dale Roark and Mike Brunolli, division fellows at Brooktree Corp., received the award for technical excellence. George Kaler, senior staff specialist at General Dynamics, received the award for the advancement of computer graphics. James Foley, professor at the Georgia Institute of Technology, received the academic award. Michael McGrath, director of the U.S. CALS program, received the award for the advancement of computer graphics standards.

NCGA's MicroCADD section presented its application award to Ruth

Ann Fraser, clinical research coordinator at the Headache Research Foundation, Faulkner Hospital, Boston; its academic award to Frank Miller, assistant professor at MIT; and its achievement award to Martin Newell, president and CEO of Ashlar.

The future?

At the session "The Future of PC Graphics: A View From the Top," J. Paul Grayson, CEO of Micrografx, went beyond many of the other panelists, who mainly talked about the future of their companies' product lines. Grayson said that soon full-motion video will be common on portable PCs. The next step will be cellular phones built into portable PCs, then voice recognition. Grayson said that someday a child will be given such a PC before kindergarten and will grow up with it, using it as an "intelligence accelerator."

And in the more immediate future, next year's NCGA conference will be in Anaheim, California, March 9 to 12.

CAD software helps reconstruct ancient abbey

Computer-aided design software normally used in the design and construction of processing plants and power stations is being used to help archaeologists reconstruct a 3D model of Furness Abbey in Cumbria, UK. Lancaster University Archaeological Unit has been carrying out a survey of the ruined Cistercian abbey on behalf of English Heritage. Along with North Cheshire College, they have used Cambridge-based CAD Centre's Plant Design and Management System to construct the 3D model from 2D stone-by-stone drawings made using photogrammetric measurement techniques.

The North Cheshire College group has stored the archaeologists' information in the PDMS database on Silicon Graphics 4D workstations. The model itself is built from predefined primitives ranging from an individual stone to complete arches, windows, vaults, roofs, and parapets. Data from the PDMS can be transferred into Enhanced Visualization System (EVS) software, an enhanced version of CAD Centre's Review software, for viewing of photorealistic, color-shaded 3D images. Viewers can examine the images from any angle and even "walk through" scenes.

According to Jason Wood, assistant director of the Lancaster University group, the visualization not only allows the archaeologists to get a feel of how the abbey would have looked and functioned, it enables them to explain the complexities of their work to people outside the field in a meaningful way.

Companies adopt emerging MPEG CD-I standard

Matsushita, Philips, and Sony have agreed to adopt the Motion Picture Experts Group (MPEG) standard for full motion video as the FMV extension of compact disc-interactive (CD-I). They have also established the CD-I Consortium Japan.

The announcement came from Gaston Bastiaens, director of Philips Interactive Media Systems, at the Sixth Annual International Conference and Exposition on Multimedia and CD-ROM in San Jose, California. He said

the consortium will create a market environment for CD-I in Japan and that similar initiatives are taking place in Europe and the United States.

Planned activities include studying the CD-I format and examining the compatibility of CD-I systems on both hardware and software, discussing the software production procedure and standardizing the user interface, exchanging information between the hardware and software industries, and promoting CD-I systems.

Kodak opens Center for Creative Imaging for artists, designers, publishers

The Professional Photography Division of Eastman Kodak has opened the Center for Creative Imaging in Camden, Maine. Originally conceived as a training site for Kodak's sales force, the center evolved into an educational and creative facility for photographers, graphic designers, art directors, computer publishers, and image composition specialists, according to Raymond DeMoulin, Kodak vice president and general manager of PDD.

DeMoulin said the company recognized a need for "an environment where professionals can experiment with constantly evolving technology." At the center, "they can master systems that combine conventional photography with electronic imaging."

The center's 50 computer work-

stations include Apple computers and Adobe software, plus equipment from other manufacturers. Teaching and conversion units networked throughout the building are tailored for specific applications with access to Kodak's Prophecy and Premier systems, film scanners, and thermal printing devices. With this equipment, participants reportedly can enhance images, perform digital composition, and expedite the integration and production of text and design.

The renovated foundry housing the center will also contain a photo gallery, cafe, classrooms, and conference rooms. The center will sponsor courses taught by a rotating faculty, events, and conferences. Joan Rosenberg is director of operations.

Said DeMoulin, "the center will be a

place where professional photographers will teach and learn, exchange ideas, and creatively advance the applications of electronic imaging."

Correction

In the May 1991 issue of *CG&A*, "Notes on Multisensory Visualization Systems" was published minus the photo credits for three of the figures. Figures 1 and 2 (page 18) ran courtesy of Scott Fisher (© 1990) and Figure 5 (page 19) courtesy of Beshers and Feiner, Columbia University (©1990). The author regrets the omissions.

IEEE Computer Graphics and applications

IEEE CG&A seeks new logo

IEEE Computer Graphics and Applications is holding a contest to find a new computer-generated logo (the graphic presentation of the magazine's name as it appears on the cover). *CG&A's* first ever computer graphics contest challenges computer graphics professionals and students to show off their skills and the capabilities of their art, design, drawing, or publishing software.

How to enter

To enter, design an attractive, distinctive logo of the words "IEEE Computer Graphics and Applications." To be consistent with the magazine's design, use a sans serif font and focus on the words "Computer Graphics."

All entries must be in black and white. The final design must fit within a space $8\frac{1}{8} \times 2$ inches (or 49×12 picas).

The month and year of the issue will appear somewhere near the logo, so if you want to, integrate them into your design.

Send a laser-printed proof, a disk of your work, and the official entry form (or photocopy). The disks can be PC (5¼-inch) or Apple Macintosh (3½-inch) disks. The electronic version must be in any of the following formats: EPS (PC or Mac), TIFF (PC or Mac), Sun Raster (on Mac disk), .PCX (PC), .IMG (PC), MacPaint (Mac), MacDraw (Mac), Pict 2 (Mac), Aldus Freehand (Mac), Corel Draw (PC), DXF (AutoCAD), IBM PIF, or HPGL plotter format. (Note: Postscript is not necessarily the same as EPS; we require EPS.)

Prizes

The winner receives \$100 and a subscription to *CG&A*. Up to three finalists also receive a subscription to *CG&A*, and all finalists will be published in a January 1992 *About the Cover* story. The story will focus on the finalists' winning designs as well as their other artwork or research.

Deadline

All submissions and official entry forms must be received by **October 15, 1991**. Send entries to *CG&A* Contest, 10662 Los Vaqueros Circle, Los Alamitos, CA 90720-1264. For more information contact Michael Haggerty at (714) 821-8380, fax (714) 821-4010.

IEEE Computer Graphics and applications

Official Contest Entry Form

Name

Address

City, State, Country

Phone Number

Fax Number

Software Used

Hardware Used

If selected as the winner, I transfer all rights for the use of this logo to IEEE.

Entrant's Signature

Date

Send a laser-printed proof, a disk of your work, and this official entry form (or photocopy) by October 15, 1991 to *CG&A* Contest, 10662 Los Vaqueros Circle, Los Alamitos, CA 90720-1264.

Book Reviews

Editor: Nick England

The best book on PHIGS

Jan Hardenbergh
Stardent Computer

A Practical Introduction to PHIGS and PHIGS-PLUS by T.L.J. Howard, W.T. Hewitt, R.J. Hubbold, and K.M. Wyrwas. (Addison-Wesley, Reading, Mass., 1991, 339 pp., hardcover, \$38.75)

It would be hard for any new book on PHIGS to *not* be the best book on PHIGS (Programmers' Hierarchical Interactive Graphics System), as there are no books on PHIGS available besides vendor documentation. But this book goes way beyond being an improvement on nothing. It is, as it claims, a practical (and thorough) introduction to PHIGS.

What is best about this book is that it's not intimidating. The authors start with an easy-to-follow "whirlwind tour" of PHIGS to get your feet wet. This introduction begins with a totally trivial example—drawing a line. Only after this example do the authors start to systematically introduce primitives, structures, and so forth.

The authors include plenty of code fragments and examples in C. Five small and two large examples give the user a good basis for developing PHIGS programs. The C binding matches last

year's standard—before underscores. Following the old standard matches the current PHIGS products, but those of us on the cutting edge need to add our own underscores. For example, "psethlhsrmode()" becomes "pset_hlhrs_mode()"—not a big deal. Nevertheless, the underscore-lacking examples are well thought out and developed along the way, and Appendixes A and B are nontrivial examples of how to do something useful in PHIGS. The first is a viewing example and the second shows what you can do with lighting.

A Practical Introduction to PHIGS and PHIGS-PLUS (or *PIP*) has 34 color plates showing variations of lighting and shading options. These include a few application screens and the standard progression from a parallel projection wireframe image to the same image in perspective with Gouraud shading and specular highlights.

Diagrams are a strong part of this book. It has all of the good diagrams that you expect—viewing frustum, deferral update flow chart, reflectance angles, a bicycle structure hierarchy chart, and so forth. It also has many new ones: attribute source flag (ASF) decision

chart, good structure-edit diagrams, and even archive-conflict resolution diagrams.

Also included in the book are chapters on output, creating and editing a model, 3D viewing, and workstations. In some of the final chapters, the authors cover structure elements for rendering control (name sets, filter, and so forth), PHIGS-PLUS graphical output, and the PHIGS-PLUS rendering pipeline.

In addition to the already mentioned example programs, the appendixes include discussions of color models and using PHIGS with Fortran as well as a summary of functions, elements, and errors. The appendix covering coordinate transforms is great, covering scaling about an arbitrary point in space clearly. This can be extended to rotation. Also included is an annotated bibliography delineating the useful papers that have been published on PHIGS recently.

The glossary seems incomplete, but this is not a reference book on all of the details and issues of PHIGS. Nor is it a read-once-and-toss tutorial. This book gives you good ideas about how to use any particular aspect of PHIGS. The authors give you enough detail to get you started, but keep things simple enough so that you can find what you want easily. This book will need to be supplemented with a good subroutine call reference like the one provided with the PEX (PHIGS Extension to X11) Sample Implementation, which should be available with the X11R5 release later this year.

But, as much as I like this book, we can always hope for a better one. The authors could have integrated both pieces of PHIGS (PHIGS 89 and PHIGS-PLUS, the International Organization for Standardization's version of the ad hoc PHIGS+ proposal). The authors could have used a newer draft of

CG&A seeks book reviewers

To help readers select worthwhile books from among the many published each year in computer graphics, *IEEE Computer Graphics and Applications* publishes book reviews (typically running from 800 to 1200 words in length). This would not be possible without the efforts of the many reviewers who volunteer their time and expertise.

We would like to add to our list of

reviewers. Interested individuals are invited to send their name, address, and areas of expertise to Nick England, CG&A Book Reviews Editor, Sun Microsystems, PO Box 13447, Research Triangle Park, NC 27709. He can also be reached via e-mail at nengland@east.sun.com. Books can be sent to England at Sun Microsystems, 2000 Aerial Center, Morrisville, NC 27560.

ADD YOUR IMAGE TO OURS

FLOPPY BOOK

*from Iterated Systems, Inc.
in cooperation with
Jones and Bartlett Publishers*

A new concept in authoring and publishing in the PC environment—a book on a single 1.4 megabyte diskette consisting of 100 pages of text and full-color pictures. Readable on a PC with VGA graphics and an 80286 or higher processor, the P.OEM PC book combines interactive control with stunning color and modular text displays. The demo book provides a model for users of this new technology, allowing the storage of a photographic 24-bit true-color image on only 10K of disk space.

Add your own image at SIGGRAPH when you purchase a copy of the Floppy book. Bring your favorite color photo to the Jones and Bartlett display (Booth #2338) to get your personalized P.OEM PC BOOK.

3.5" Disk, \$24.95, Aug. 1991
0-86720-222-X

...Learn more about the fractal transform and the P.OEM system in a forthcoming book by Michael Barnsley/Iterated Systems. In this follow-up to the popular FRACTALS EVERYWHERE, Michael Barnsley/Iterated Systems explain the workings of the fractal transform and describe the P.OEM system which exploits this new technology for data compression.

MULTIPROCESSOR METHODS FOR COMPUTER GRAPHICS RENDERING

Scott R. Whitman

This book is a thorough investigation into the problem of using a massively parallel computer to render three-dimensional computer graphics scenes into images.

ca. 300 pp., \$39.95, Oct. 1991, 0-86720-229-7

COMPUTER GRAPHICS FILE FORMATS: A Taxonomy and General Reference Guide

Wayne Brown, Sam Reynolds

This text provides a cohesive reference to graphical file formats. The taxonomy developed categorizes the file formats in such a way that the reader can clearly understand the types of data that a file format can contain, how the format organizes the data and whether the data can be easily shared between or transformed into other file formats.

ca. 350 pp., \$39.95, 1992
0-86720-235-1

GEOMETRIC CONCEPTS FOR GEOMETRIC DESIGN

*Wolfgang Boehm
Hartmut Prautzsch*

This unique and lively survey of geometric concepts is written with an eye on applications and implementation in geometric design and computer graphics.

ca. 350 pp., \$39.95, Dec. 1991, 0-86720-226-2

ITERATED FUNCTION SYSTEMS AND THE PROPAGATION OF ROUNDING ERRORS

Mario Peruggia

The author investigates the propagation of rounding errors in the context of a novel method for encoding a wide class of digitized color images using the identification of the image with a unique stationary Markov chain. The author describes how IFS theory can be applied to produce animated motion pictures.

ca. 250 pp., \$29.95, Dec. 1991, 0-86720-236-X

A PHYSICAL APPROACH TO COLOR IMAGE UNDERSTANDING

Gudrun Klinker

The author uses a physical, intrinsic reflection model to analyze the effects of shading and highlights on dielectric, non-uniform, opaque materials. When combined with a camera model, this description of physically possible color changes on one object can be used to guide an algorithm in interpreting color images.

ca. 200 pp., \$29.95, Sept. 1991, 0-86720-237-8

WAVELETS

edited by G. Beylkin,

R. Coifman, I. Daubechies,

S. Mallat, Y. Meyer, L. Raphael

and M. B. Ruskai

This text provides state-of-the-art surveys by leading researchers in many of the applications areas where wavelets are making an impact.

ca. 400 pp., \$48.75, Sept. 1991, 0-86720-225-4

MOBILE ROBOTS: Inspiration to Implementation

Anita Flynn, Joseph Jones

This unique text is a complete guide to building mobile robots. It contains everything the reader needs to know to build a robot in a progression of design projects from the simplest form to more complex artificial intelligence systems—and where to go to find the parts!

ca. 250 pp., \$29.95, Dec. 1991
0-86720-223-8

the "C" binding. Although the decision to use the one that is in the field was good, it won't be in the field for long. But these are small flaws. I am hoping to find two more books on PHIGS at this year's Siggraph from Prentice Hall (Valerie Clark) and Wiley & Sons (Hopgood and Duce). After Siggraph I expect to see at least five more books on PHIGS, and I expect a grand total of at least eight by Siggraph 1992.

I'm sure some of these other books will provide much more detail on all the facets of PHIGS. *The Practical Introduction to PHIGS and PHIGS-PLUS* does not try to give in-depth explanations of every detail. What it does give is enough information for you to try the specific options PHIGS presents. This allows the authors to deliver a book that covers all of PHIGS but is still a compact, to-the-point package. I appreciate that, and I think many other readers will too.

This book has been criticized for light coverage of deferral and modification modes or quick update methods, important features of PHIGS that allow interactive applications to modify part of the graphics without redisplaying the whole picture. This is a fair criticism; the examples will not work on all PHIGS implementations without adding some update-workstation calls. But of course no PHIGS examples will work on all PHIGS implementations until all of the implementations use the final standard C bindings.

This book has been criticized for its limited coverage of NURBS (nonuniform rational B-splines) and other PHIGS-PLUS primitives. The authors cover enough of the bread and butter PHIGS-PLUS primitives (such as triangle strip and quadrilateral mesh), lighting, and shading so that I can honestly state that they cover PHIGS-PLUS. They write enough about NURBS to allow you to call the functions, then they refer you to good books that cover curves and surfaces in detail. In light of the fact that the PHIGS-PLUS standard is not final, I think the authors made good trade-offs between what they covered and what they did not.

This book is not the standard warmed over or breezy tour of PHIGS as was *Understanding PHIGS* (Maxine Brown and Michael Heck, Megatek, 1985) and the chapter in Foley, van Dam, Feiner, and Hughes (*Computer Graphics: Principles and Practice*, Addison-Wesley, 1990). This book is PHIGS explained as you need it by people who have obviously explained it before, many times.

Reader Service Number 20

IEEE Computer Graphics & Applications

Jones and Bartlett Publishers

20 Park Plaza • Boston • MA 02116 • 617-482-3900 • 800-832-0034



New Products

Editors: Carl Machover and John C. Dill

HARDWARE

Vistra 800 brings visualization to the desktop

Stardent Computer has announced the Vistra 800 Series of desktop visualization systems based on Intel's 40-MHz 860 microprocessor, which has built-in graphics processing capabilities. The three models in the series run Stardent's AVS visualization software.

According to the company, the products offer integrated graphics and rendering capabilities, including single and multipass true transparency, 2D and 3D texture mapping with integrated lighting, and global antialiasing. Other built-in capabilities include depth cueing and fog; backface culling and automatic normal flipping; flat, Gouraud, Phong, and quadratic shading; ambient, directional, point, and spot lighting models from 16 different colored light sources; and both diffuse and specular highlights.

The Vistra 800 is a 16-bit true color

system with RGB control. It reportedly renders 33,000 polygons and 260,000 3D vectors per second. The Vistra 800e includes an extended frame-buffer architecture for 24-bit true color and reportedly renders 32,000 polygons and 260,000 3D vectors per second. The Vistra 800ex includes two additional 860 processors and 16 Mbytes of additional virtual pixel map memory. According to the company, this provides 90,000 polygons and 500,000 3D vectors per second in 24-bit color.

All three models come with a 3.5-inch 1.44-Mbyte floppy disk drive and a 3.5-inch drive with either 200 or 400 Mbytes of memory. A SCSI interface supports external storage devices and peripherals. Prices range from \$20,995 to \$59,015.

Reader Service Number 30

Video Toaster stands alone

Newtek now offers a stand-alone version of its Video Toaster system for broadcast-quality video production. The desktop system includes custom chips that perform real-time video manipulation, plus software for video switching, digital video effects, character generation, color processing, 3D animation, and paint graphics.

The Video Toaster stand-alone system includes an M68000 CPU (optional 68030 or 68040 coprocessor), two-button mouse and 94-key keyboard, 880-Kbyte micro floppy drive, 52-Mbyte SCSI hard drive, 5 Mbytes of RAM (expandable to 9 Mbytes), custom graphic coprocessors, four-voice stereo audio output, and a windowing multitasking operating system.

The Video Toaster costs \$3,995.

Reader Service Number 32

Tektronix claims first PEX terminal

Tektronix has announced what it calls the first PEX (PHIGS Extensions to X) terminal, the Tek Xpress XP29P, for the development of distributed 2D and 3D applications. The terminal is based on the MIT X Consortium's initial PEX5R1 release and supports Version 5.0P of the PEX protocol.

The XP29P has a multiprocessor architecture built around Motorola's 68030 for network traffic and Texas Instruments' 34020 graphics coprocessor and 34082 floating-point processor for image rendering. It comes standard with 7 Mbytes of memory expandable to 39 Mbytes, 128 Kbytes of ROM, keyboard, and three-button mouse, plus support for a tablet and trackball. It also has a 19-inch Panasonic monitor with 1,280 × 1,024 resolution and 256 colors displayable from a palette of 16.7 million.

Prices for the XP29P start at \$6,995.

Reader Service Number 31



Tektronix' Tek Xpress XP29P terminal supports PEX for development of 2D and 3D applications.

DEC RISC/Unix workstations feature upgradeable CPU card

Digital Equipment has announced midrange Unix-based RISC workstations, the DECstation 5000/100 Series, with an unusual upgrade feature—a replaceable central processing unit/floating-point unit (CPU/FPU) daughter card. The two series models come with a choice of five different graphics configurations and the OSF/Motif graphical user interface. According to the company, the Model 120 and 125 workstations are designed for computation-intensive 2D and 3D graphics applications. Prices start at \$6,495.

DECstation 5000/100 models come with a R3000A CPU/R3010 FPU chip

set from Mips Computer Systems on a 3 × 5-inch removable daughter card. Model 120 runs at 20 MHz and Model 125 at 25 MHz. Options include the entry-level monochrome MX, color and grey-scale graphics CX, raster graphics PX, 3D modeling and imaging graphics PXG, and 3D modeling and visualization PXG Turbo.

PXG and PXG Turbo systems come with Personal Visualizer software from Wavefront Technologies. Nine pre-configured systems available include the RRD42 CD-ROM drive, a 5.25-inch half-height device with a capacity of 600 Mbytes.

Reader Service Number 33

Skywriter family assists real-time simulation

Silicon Graphics targets real-time simulation with its Skywriter family of mid-range visual systems. The company claims that the systems' combination of a parallel computing architecture with a realistic scene generator enables users to integrate a complete simulator with a single system.

Skywriter systems come in 62-inch racks with two graphics pipelines driving one or two independent output channels. Configurations include two or four parallel processors and up to 256 Mbytes of main memory and 38 Gbytes of secondary storage. Graphics options reportedly enable a single system to provide up to eight output channels, composite video output, and video input. Prices start at \$209,900.

Image generation features include photo-derived texture mapping, anti-aliasing, atmospheric effects, and special effects for sensor simulation. The company claims the systems can display up to 5,000 polygons per frame at 30 Hz with full texture mapping and filtering.

Skywriter's real-time host computer is based on Silicon Graphics' Powerpath parallel architecture. Skywriter incorporates the React kernel for real-time and parallel processing and is controlled by Irix, a version of Unix. The Skywriter systems are compatible with Iris workstations and servers.

Reader Service Number 35

Image processor integrates vector and raster data

Intergraph has announced the Imagestation 6187 image processing workstation. The system includes an integrated parallel architecture that reportedly permits simultaneous viewing and manipulation of vector and raster data on a single screen. Based on Intergraph's Series 6000 RISC-based Unix workstations and Edge II graphics subsystem, the 6187 adds image processing technology from Visual Information Technologies.

According to the company, the 6187 has 24-bit color and 8-bit overlay dis-

play capabilities and operates at 300 million operations per second and 18 million floating-point operations per second. The Edge II graphics subsystem provides vector processing and polygon processing. The modified Edge II frame buffer produces a simultaneous vector-over-raster display on a single screen.

According to the company, applications that require stereo technology delivered with the system will be available in 1991.

Prices start at \$107,000.

Reader Service Number 34

SOFTWARE

Generic Software offers limited-time upgrade of Generic CADD for the Macintosh for \$99

Generic Software offers a \$99 limited-time upgrade to existing Generic CADD users with MacDraw, MacDraft, Claris CAD, or other Macintosh draw or CAD programs. The package usually retails for \$595. The offer is good to September 30, 1991.

The latest version of Generic CADD for the Macintosh includes more than 100 new drafting and editing features, according to the company. These include snap previewing, which allows users to see the results of snap commands before placing them; automated associative dimensioning, in which dimensions change automatically to reflect changes in the drawing; and

symbol preview, which allows users to preview symbol images before placing them (from Generic's libraries or the users' own symbol library). The software comes as a design and drafting system, including the DXF translator, with 14 optional symbol libraries available.

Generic CADD for the Macintosh operates in System 6.0.3 or later. For color, users need System 6.0.5 or later and Quick Draw (included). The software requires a minimum of a Mac Plus or later with 1 Mbyte of memory. The company recommends a hard drive and math coprocessor, plus 2 Mbytes or more of RAM.

Reader Service Number 40

Computervision markets Design View through VARs

Computervision acquired the assets and hired the Design View software development team of Premise, Inc. in Cambridge, Mass. Computervision will offer Design View parametric design software through its network of value-added resellers.

The company says that engineers using Design View can simply sketch a basic shape, focusing on the relationships between mechanical components rather than their sizes. When the user later specifies dimensions, the software automatically reshapes the geometry.

Design View version 2.0 lists for \$895.

Reader Service Number 41

Software simulates electromagnetic fields

Ansoft has announced the availability of its Maxwell 3D Field Simulator software. According to the company, the CAE software provides sophisticated finite element analysis techniques while eliminating the need for user interaction with the FEA process. It also incorporates algorithms that reduce memory and processor hardware requirements.

The Maxwell 3D Field Simulator is available for Unix workstations such as Sun, Hewlett-Packard, and IBM platforms. It can be licensed annually for \$18,000 or purchased for \$45,000 with a perpetual license.

Reader Service Number 42

Images created under Windows 3

Rix Softworks claims that its Winrix software permits true-color image creation and editing with 32,000 (15-bit) or 16.8 million (24-bit) colors on a VGA monitor. The company says it designed the product as a Windows 3 application.

Winrix uses pull-down menus for major menu selection and pop-up menus for secondary menu selection. It reportedly offers multiple document interface and dynamic document exchange, plus image-editing tools and color support for vendor-supplied drivers up to 24-bit color.

Winrix retails for \$495. Registered Colorix users can upgrade for \$100 to \$200 depending on their version of that package.

Reader Service Number 43

Kodak XL700 printers link to Unix workstations

Eastman Kodak has announced that Z Microsystems' Imagejet.PS software now supports the Kodak XL 7700 digital continuous tone printer. The company says Unix workstations running graphics-intensive applications can output high-resolution full color prints or transparencies up to 11 x 11 inches from both Postscript and raster image files through the SCSI port.

Imagejet software includes modules allowing the selective capture of raster images from screens produced by an X Windows, Sunview, or Open Windows interface. Other utilities permit clipping, rotation, scaling, cut-and-paste, and other operations before printing.

Reader Service Number 44

New name for quality improvement software

Catalyst has announced a new name, new pricing, and a new release of its manufacturing problem-solving and quality improvement software. The Macintosh-based software, called Catalyst/DOE (formerly Catalyst/RPE), reportedly simplifies design of experiments. According to the company, it is a graphics tool that allows engineers to use statistical methods to solve manufacturing problems without advanced statistical training.

Catalyst/DOE Release 1.3 has a num-

ber of new features, including clipboard support, an enhanced graphical interface, quadratic modeling, and model term selection. A license for one workstation costs \$1,995, with an annual renewal fee of \$995. The license includes unlimited telephone technical support and automatic software upgrades. Site licensing starts at \$25,000, with a \$12,500 renewal fee. Full-site licenses include on-site training, technical support, and software upgrades.

Reader Service Number 45

VAPS helps commercial companies develop GUIs

Virtual Prototypes now offers its 2.0 version of VAPS (Virtual Applications Programming System) in the commercial marketplace. VAPS has previously been used in aerospace and related industries for the design and real-time animation of graphical user interfaces.

The company says that VAPS uses a strictly pictorial approach and can transport display code to the system that will run the application. With the pictorial programming approach, objects are assigned visual and behavioral characteris-

tics and are then integrated with application data. VAPS shares the workstation with the user application or can communicate with it over Ethernet. GUIs built with VAPS can be distributed over several workstations or driven from several applications residing on a network, according to the company.

Prices start at \$15,000. VAPS supports Unix-based workstations such as those in the IBM RS/6000, Sun Sparcstation, SGI, and HP Apollo series.

Reader Service Number 46

Euclid-IS expands with adaptive modeling

Matra Datavision has announced Adaptive Modeling software for its Euclid-IS system. The company claims the new package provides extensive capabilities in the areas of parametric modeling and dimension-driven feature editing. The Parametric Geometry module is a \$2,000 option to Euclid-IS, while the Feature Editor module is free.

The company calls Euclid-IS a CAE/CAD/CAM system that uses solid and surface modeling technology and a distributed object-oriented database to integrate and manage product design, analysis, documentation, and manufacturing activities. With Adaptive Modeling modules users can create and modify solid models using parametric techniques and edit features interactively. Integration with the Euclid-IS system means that changes are propagated throughout the system.

According to the company, Adaptive Modeling allows users to decide what to parameterize and what to constrain. Users can combine free geometry with constrained geometry, modifying parts freely regardless of constraints defined earlier.

Reader Service Number 47

Classified Ad

CLASSIFIED RATES: \$5.00 per line, \$50 minimum charge (up to ten lines). Average five typeset words per line, eight lines per column inch. Send copy at least one month prior to publication to: Marian Tibayan, Classified Advertising, *IEEE Computer Graphics & Applications Magazine*, 10662 Los Vaqueros Circle, Los Alamitos, CA 90720-1264; (714) 821-8380; fax (714) 821-4010.

TMI High Technology Sales

CAD MONITORS, GRAPHICS DIGITIZER TABLETS. Thermal printers, pencil/pen plotters. Graphic controller cards: Number Nine, Vermont Microsystems, etc. Monitors: Nanao, Taxan, etc. Call/fax for information or prices. This month's Trinitron flat-screen monitor specials: Seiko CM-2050 20" 1024x768 includes Orchid Pro II video card \$1,925. Taxan Ultravision 1095 20" 1600x1280 \$2,895. TMI High Technology Sales (714) 639-6085.

Advertiser/Product Index

July 1991

Advertisers

Academic Press, Inc.	Cover II
ICCIM '91	Cover IV
IEEE Computer Society Membership	23
IEEE Computer Society Press	66, 74, 80
Jones and Bartlett Publishers	92
Science Accessories Corp.	4
Classified Advertising	95

For display advertising, contact:

Northwest Region: Janet Engelbrecht; Tel: (415) 905-0260; Fax: (415) 896-1512.

Eastern Region: Georgette Boone; Tel: (415) 905-0260; Fax: (415) 896-1512.

Southwest/Southeast Region: Kenneth Smitley; Tel: (415) 905-0260; Fax: (415) 896-1512.

Recruitment and Classified Advertising: D. Rodney Brooks;
Tel: (415) 905-0260; Fax: (415) 896-1512.

Director of Sales: Randall L. Stickrod, 544 Second St., Suite 200, San Francisco, CA 94107; Tel: (415) 905-0260; Fax: (415) 896-1512.

Publications Office: *IEEE Computer Graphics & Applications*, 10662 Los Vaqueros Circle, PO Box 3014, Los Alamitos, CA 90720; (714) 821-8380; fax: (714) 821-4010.

Products

	RS #	Page #
Academic Press, Inc.	1	C.II
Addison Wesley	20	91
Ansoft	42	95
Catalyst	45	95
Computervision	41	94
Digital Equipment	33	94
Eastman Kodak	44	95
Generic Software	40	94
Intergraph	34	94
Jones and Bartlett Pub.	3	92
Matra Datavision	47	95
Newtek	32	93
Rix Softworks	43	95
Science Accessories	2	4
Silicon Graphics	35	94
Stardent Computer	30	93
Tektronix	31	93
Virtual Prototypes	46	95

IEEE Computer Graphics and applications

July 1991 issue (card void after January 1991)

Name _____

Title _____

Company _____

Address _____

City _____

State _____ ZIP _____

Country _____

Phone (_____) _____

Reader Interest

(Circle what you liked, add comments on the back.)

- A** Ray Tracing
- B** Scouting
- C** Numbers
- D** Ancient Culture
- E** Development
- F** Learning
- G** Primitives
- H** Demise
- L** Letters
- M** About the Cover
- N** Displays on Display
- O** Visualization
- P** Editors' Intro.
- Q** Jim Blinn's Corner
- R** Selective Update
- T** Book Reviews
- V** New Products
- Z** Final Focus

Product Information

(Circle the numbers to receive product information.)

- | | | | | | |
|----|----|----|----|-----|-----|
| 1 | 23 | 45 | 67 | 89 | 111 |
| 2 | 24 | 46 | 68 | 90 | 112 |
| 3 | 25 | 47 | 69 | 91 | 113 |
| 4 | 26 | 48 | 70 | 92 | 114 |
| 5 | 27 | 49 | 71 | 93 | 115 |
| 6 | 28 | 50 | 72 | 94 | 116 |
| 7 | 29 | 51 | 73 | 95 | 117 |
| 8 | 30 | 52 | 74 | 96 | 118 |
| 9 | 31 | 53 | 75 | 97 | 119 |
| 10 | 32 | 54 | 76 | 98 | 120 |
| 11 | 33 | 55 | 77 | 99 | 121 |
| 12 | 34 | 56 | 78 | 100 | 122 |
| 13 | 35 | 57 | 79 | 101 | 123 |
| 14 | 36 | 58 | 80 | 102 | 124 |
| 15 | 37 | 59 | 81 | 103 | 125 |
| 16 | 38 | 60 | 82 | 104 | 126 |
| 17 | 39 | 61 | 83 | 105 | 127 |
| 18 | 40 | 62 | 84 | 106 | 128 |
| 19 | 41 | 63 | 85 | 107 | 129 |
| 20 | 42 | 64 | 86 | 108 | 130 |
| 21 | 43 | 65 | 87 | 109 | 131 |
| 22 | 44 | 66 | 88 | 110 | 132 |

1

IEEE Computer Graphics and applications

July 1991 issue (card void after January 1991)

Name _____

Title _____

Company _____

Address _____

City _____

State _____ ZIP _____

Country _____

Phone (_____) _____

Reader Interest

(Circle what you liked, add comments on the back.)

- A** Ray Tracing
- B** Scouting
- C** Numbers
- D** Ancient Culture
- E** Development
- F** Learning
- G** Primitives
- H** Demise
- L** Letters
- M** About the Cover
- N** Displays on Display
- O** Visualization
- P** Editors' Intro.
- Q** Jim Blinn's Corner
- R** Selective Update
- T** Book Reviews
- V** New Products
- Z** Final Focus

Product Information

(Circle the numbers to receive product information.)

- | | | | | | |
|----|----|----|----|-----|-----|
| 1 | 23 | 45 | 67 | 89 | 111 |
| 2 | 24 | 46 | 68 | 90 | 112 |
| 3 | 25 | 47 | 69 | 91 | 113 |
| 4 | 26 | 48 | 70 | 92 | 114 |
| 5 | 27 | 49 | 71 | 93 | 115 |
| 6 | 28 | 50 | 72 | 94 | 116 |
| 7 | 29 | 51 | 73 | 95 | 117 |
| 8 | 30 | 52 | 74 | 96 | 118 |
| 9 | 31 | 53 | 75 | 97 | 119 |
| 10 | 32 | 54 | 76 | 98 | 120 |
| 11 | 33 | 55 | 77 | 99 | 121 |
| 12 | 34 | 56 | 78 | 100 | 122 |
| 13 | 35 | 57 | 79 | 101 | 123 |
| 14 | 36 | 58 | 80 | 102 | 124 |
| 15 | 37 | 59 | 81 | 103 | 125 |
| 16 | 38 | 60 | 82 | 104 | 126 |
| 17 | 39 | 61 | 83 | 105 | 127 |
| 18 | 40 | 62 | 84 | 106 | 128 |
| 19 | 41 | 63 | 85 | 107 | 129 |
| 20 | 42 | 64 | 86 | 108 | 130 |
| 21 | 43 | 65 | 87 | 109 | 131 |
| 22 | 44 | 66 | 88 | 110 | 132 |

2

CG&A

IEEE CG&A . . . because graphics shouldn't be all work and no display

☐ YES, sign me up!

If you are a member of the Computer Society or any other IEEE society, pay the member rate of only \$24 for a year's subscription (six issues).

Society: _____

IEEE membership no.: _____

Society members: Subscriptions are annualized. For orders submitted March through August, pay half the full-year rate (\$12) for three bimonthly issues.

Full Signature _____ Date _____

Name _____

Street _____

City _____

State/Country _____ ZIP/Postal Code _____

☐ YES, sign me up!

If you are a member of ACM, CGS, EACG, IEEE (but not a member of an IEEE society), NCGA, SIGGRAPH, or other professional society, pay the sister-society rate of only \$39 for a year's subscription (six issues).

Organization: _____

Membership no.: _____

☐ Payment enclosed

☐ Charge to ☐ Visa ☐ MasterCard ☐ American Express
Charge-card number _____

Expiration date _____

Month _____ Year _____

Prices valid through 12/31/91

7/91 CG&A

Charge orders also taken by phone:
(714) 821-8380 8 a.m. to 5 p.m. Pacific time
Circulation Dept.
PO Box 3014
Los Alamitos, CA 90720-1264

Editorial comments

I liked: _____

I disliked: _____

I would like to see: _____

For reader-service inquiries,
see other side.



PLACE
POSTAGE
HERE

This address is for reader-service cards only.

IEEE CG&A

Reader Service Inquiries
PO Box 16508
North Hollywood, CA 91615-6508
USA



Editorial comments

I liked: _____

I disliked: _____

I would like to see: _____

For reader-service inquiries,
see other side.

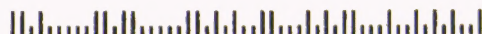


PLACE
POSTAGE
HERE

This address is for reader-service cards only.

IEEE CG&A

Reader Service Inquiries
PO Box 16508
North Hollywood, CA 91615-6508
USA



NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

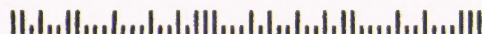
BUSINESS REPLY MAIL

FIRST CLASS PERMIT NO. 38 LOS ALAMITOS, CA

POSTAGE WILL BE PAID BY ADDRESSEE

IEEE COMPUTER SOCIETY

PO Box 3014
Los Alamitos, CA 90720-9804
USA



Final Focus



©1991 Pete Segal

Mona's Suburban Cafe

Pete Segal

AT&T Pixel Machines

Holmdel, New Jersey



INTERNATIONAL CONFERENCE ON COMPUTER INTEGRATED MANUFACTURING

CONFERENCE & EXHIBITION 2 - 4 October 1991
TUTORIALS 30 Sept - 1 Oct 1991

Hyatt Regency Singapore

The First International Conference on Computer Integrated Manufacturing will be held in Singapore on 2 to 4 October 1991. The conference is organised by GINTIC Institute of CIM, Nanyang Technological University, Singapore. The theme of this conference is *Manufacturing Enterprises of the 21st Century*. In conjunction with the conference, there will be tutorial sessions (30 Sept - 1 Oct 1991) and a major exhibition at the conference site.

The first call for papers attracted a total of 216 practical and academic exponents from 28 countries. The papers will be reviewed by an International Advisory Panel, comprising experts in the field of CIM from 15 different countries, and the National Technical Advisory Panel.

Chairman, Organising Committee
A/Prof Robert Gay

Chairman, National Technical Advisory Panel
Dr Lim Beng Siong

Chairman, Publicity & Sponsorship Sub-committee
Mr Tay Kim Huat

Chairman, Exhibition Sub-committee
Mr Seah Hock Soon

Conference Secretariat
Associated Conventions & Exhibitions Pte Ltd

Keynote Addresses

In addition to technical sessions, there will be 10 keynote addresses and 2 invited papers by world-renown experts in their respective fields. They will address the contemporary issues relating to CIM.

Tutorial Sessions

The two-day tutorial prior to the conference will cover three broad areas, namely CIM-oriented Issues; Focus on Industrial Artificial Intelligence Applications; and Concurrent Engineering, Design and Simulation Aspects.

ICCIM '91 Exhibition

The 3-day exhibition will feature the latest state-of-the-art technology in CIM. Organisations wishing to take part in this exhibition may request for the Exhibitors Information Package from the conference secretariat to register as an exhibitor.

Technical Sessions

The topics of the technical sessions can be broadly categorised under the four main groups: computer integrated manufacturing (CIM); artificial intelligence (AI); flexible manufacturing systems (FMS); and simulation.

For more information on the ICCIM '91 conference, tutorial and exhibition, please fax the information request slip to the Conference Secretariat, or call Ms Rena Lim at Tel: (65) 732-6839.

To: Rena Lim, The Conference Secretariat
Associated Conventions and Exhibitions Pte Ltd
204 Bukit Timah Road
#04-00 Boon Liew Building
Singapore 0922
Fax: (65) 732-6309

INTERNATIONAL CONFERENCE ON COMPUTER INTEGRATED MANUFACTURING

Conference & Exhibition 2 - 4 October 1991
Tutorials 30 Sept - 1 Oct 1991
Hyatt Regency Singapore

INFORMATION REQUEST SLIP

Please send me the following items

- ☐ The Advance Programme
☐ Exhibitors Information Package
☐ Sponsorship Information Package

Prof/Dr/Mr/Mrs/Ms: _____

Designation: _____

Organisation: _____

Address: _____

Telephone: _____ Fax: _____